



# JOINT HIGHWAY RESEARCH PROJECT

IN/JHRP-79-18

RELATIONSHIP OF HIGHWAY  
DEVELOPMENT AND CITY  
DEVELOPMENT FOR  
NON-METROPOLITAN  
PLACES IN INDIANA

Hadi Yamin





Final Report

RELATIONSHIP OF HIGHWAY DEVELOPMENT AND CITY DEVELOPMENT  
FOR NON-METROPOLITAN PLACES IN INDIANA

TO: Harold L. Michael, Director  
Joint Highway Research Project

October 23, 1979

Project: C-36-32V

FROM: Robert D. Miles, Research Engineer  
Joint Highway Research Project

File: 1-4-22

Attached is the Final Report titled "Relationship of Highway Development and City Development for Non-Metropolitan Places in Indiana". It has been authored by Mr. Hadi Yamin, Graduate Instructor in Research on our staff, under the direction of Professor Robert D. Miles.

The publication reports the development of a generalized eleven variable transportation-city development path model. It was calibrated separately with 1940, 1950, 1960 and 1970 data. The cities samples were representative of non-metropolitan cities with population between 2500 and 50,000. Twenty equations (five land uses in four time frames) were developed and fifteen equations explained 73 to 98 percent of measured land use variations. The four path models representing the four decades were qualitatively different and it was not feasible to combine into an overall model for a continuous time frame. The calibrated 1970 model can be used to estimate land use changes as consequences of future planning decisions. Total effect coefficients obtained from the 1970 path model may be very useful in environment impact studies of proposed transportation or industrial projects.

The Final Report is submitted for acceptance as fulfillment of the objectives of the Study.

Respectfully submitted,



Robert D. Miles, P.E.

RDM:ms

cc: A. G. Altschaeffl  
W. L. Dolch  
R. L. Eskew  
G. D. Gibson  
W. H. Goetz  
M. J. Gutzwiller  
G. K. Hallock

D. E. Hancher  
K. R. Hoover  
J. F. McLaughlin  
R. F. Marsh  
R. D. Miles  
P. L. Owens  
G. T. Satterly

C. F. Scholer  
K. C. Sinha  
C. A. Venable  
L. E. Wood  
E. J. Yoder  
S. R. Yoder



Final Report

RELATIONSHIP OF HIGHWAY DEVELOPMENT AND CITY DEVELOPMENT  
FOR NON-METROPOLITAN PLACES IN INDIANA

by

Hadi Yamin  
Graduate Instructor in Research

Joint Highway Research Project

Project No.: C-36-32V

File No.: 1-4-22

Prepared as Part of an Investigation

Conducted by

Joint Highway Research Project  
Engineering Experiment Station  
Purdue University

in cooperation with the  
Indiana State Highway Commission

Purdue University  
West Lafayette, Indiana  
October 23, 1979

Digitized by the Internet Archive  
in 2011 with funding from  
LYRASIS members and Sloan Foundation; Indiana Department of Transportation

## ACKNOWLEDGMENTS

The author expresses sincere appreciation to Professor Robert D. Miles, his major professor, for his expert guidance and support throughout the course of this project, and for critical review of the manuscript.

The author is grateful to Professor Harold L. Michael who was his major advisor up to his passage of Ph.D. preliminary exams and approval of the proposal for this research, for his support, academic guidance, assistance in development of the research proposal, and for his critical review of transcript.

Special thanks are due to other members of the committee. Professor Virgil Anderson gave many invaluable advice and suggestions in statistical design and analysis of the project. Professor Harvey H. Marshall contributed greatly to the project with his numerous advice and suggestions, particularly with his suggestion of using path model as a means of obtaining statistical control of the study.

The author is grateful to Professor Kumares C. Sinha for his moral support and invaluable guidance throughout the project and for his critical review of the major portion of the manuscript.

The writer is grateful to Purdue University Joint Highway Research Project and Indiana State Highway Commission for financial support of this research project.





Thanks are also due the Indiana State Highway Commission and U.S. Geological Survey for the loan of required aerial photos not available in the Joint Highway Airphoto Laboratory. Thanks are due to Indiana University Geography Library for loan of 1970 place maps for Indiana cities.

Thanks go to the authors colleagues James Mekemson and Anil Bhandari for their moral support and assistance as needed. Thanks are due Messrs. Anthony Okafour, Michael Frutt, and Hossien Rostam, undergraduate civil engineering students for their help during the data collection phase. Special thanks are due Mrs. Marian Sipes for the excellent typing of the manuscript; and to Mr. Emmett Black for his expert photographic reduction of numerous tables and graphs included in the report.

Finally special thanks are extended to my wife, Parvaneh, for her encouragement, and many sacrifices during the term of this project; and to my parents without whom this endeavor would not have been possible.



## TABLE OF CONTENTS

	<u>Page</u>
LIST OF TABLES. . . . .	vi
LIST OF FIGURES . . . . .	x
ABSTRACT. . . . .	xii
CHAPTER ONE: INTRODUCTION. . . . .	1
Purpose. . . . .	3
Scope. . . . .	3
References . . . . .	4
CHAPTER TWO: PREVIOUS INVESTIGATIONS . . . . .	5
Types of Transportation Impact Studies . . . . .	5
Bypass Studies. . . . .	6
Impact of Radial Highways on Narrow Corridors . . . . .	7
Impact of Major Highway Construction on a Region. . . . .	8
Interrelationship of Transportation and Land Development. . . . .	9
Summary of Previous Transportation Impact Studies and Their Methodology . . . . .	10
Analysis of Transportation Impact Studies Using "Before and After" Methodology. . . . .	11
References . . . . .	17
CHAPTER THREE: A FRAMEWORK FOR DEVELOPING THE INTERRELATIONSHIP OF TRANSPORTATION DEVELOPMENT AND CITY DEVELOPMENT . . . . .	20
Use of Path Analysis for Verification of Relationship Between Transportation Development and City Development . . . . .	22
Total Effect of an Independent Variable on a Dependent Variable. . . . .	29
Path Model Hypothesized to Explain Relationship Between Transportation Development and City Development . . . . .	30
References . . . . .	40



## TABLE OF CONTENTS (Continued)

	<u>Page</u>
CHAPTER FOUR: DESIGN OF QUASI-EXPERIMENT . . . . .	41
Procedure to Select Sample Cities. . . . .	43
Data Collection. . . . .	66
City Location (X1). . . . .	69
Inter-City Accessibility (X2) . . . . .	69
Intra-City Accessibility (X3) . . . . .	89
Land Use Variables (X4, X8, X9, X10, X11) . . . . .	90
Census Data (X5, X6, X7). . . . .	102
Statistical Analysis and Results . . . . .	107
Calibration of the Path Models. . . . .	107
Results of 1940, 1950, 1960 and 1970 Model	
Calibrations . . . . .	114
1940 Path Model. . . . .	115
1950 Path Model. . . . .	120
1960 Path Model. . . . .	125
1970 Path Model. . . . .	130
Ability of Calibrated Path Models to Predict Land Use	
Variables. . . . .	136
Qualitative Comparison of the Four Calibrated Models. .	143
Qualitative Comparison of the Four Equations	
Predicting Residential Land Use . . . . .	144
Qualitative Comparison of the Four Equations	
Predicting Institutional Land Use . . . . .	146
Qualitative Comparison of the Four Equations	
Predicting Vacant Land Use. . . . .	147
Qualitative Comparison of the Four Equations	
Predicting Commercial Land Use. . . . .	148
Qualitative Comparison of the Four Equations	
Predicting Industrial Land Use. . . . .	148
Results of Qualitative Comparisons of the	
Four Models . . . . .	148
Total Effect of an Independent Variable on a Dependent	
Variable. . . . .	149
Procedure for Calculation of Total Effects of	
Independent Variables on the Dependent Variables	
for the General Path Model . . . . .	149
Results Obtained from Calculation of Total Effects	
of Independent Variables on Dependent Variables. .	152
References . . . . .	160
CHAPTER FIVE: CONCLUSIONS. . . . .	165
LIST OF REFERENCES. . . . .	170
APPENDICES. . . . .	176
Appendix A . . . . .	176
Appendix B . . . . .	209
Appendix C . . . . .	239



## LIST OF TABLES

<u>Table</u>		<u>Page</u>
1	Assumptions Used in Development of the Path Model Describing the Relationship of Transportation Development and City Development. . . . .	33
2	Designation and Names of Variables Used in the Transportation City Development Path Model. . . . .	38
3	List of 1940 Cities in the Sample Space - Northern Zone. . . . .	53
4	List of 1940 Cities in the Sample Space - Central Zone. . . . .	54
5	List of 1940 Cities in the Sample Space - Southern Zone. . . . .	55
6	List of 1950 Cities in the Sample Space - Northern Zone. . . . .	56
7	List of 1950 Cities in the Sample Space - Central Zone. . . . .	57
8	List of 1950 Cities in the Sample Space - Southern Zone. . . . .	58
9	List of 1960 Cities in the Sample Space - Northern Zone. . . . .	59
10	List of 1960 Cities in the Sample Space - Central Zone. . . . .	60
11	List of 1960 Cities in the Sample Space - Southern Zone. . . . .	61
12	List of 1970 Cities in the Sample Space - Northern Zone. . . . .	62
13	List of 1970 Cities in the Sample Space - Central Zone. . . . .	63
14	List of 1970 Cities in the Sample Space - Southern Zone. . . . .	64





## LIST OF TABLES (Continued)

<u>Table</u>		<u>Page</u>
15	Number of Cities in the Sample Space and Number of Cities Sampled in Each of the Cells . . . . .	65
16	Cross-Sectional Samples of Non-Metropolitan Indiana Cities for 1940 . . . . .	67
17	Cross-Sectional Samples of Non-Metropolitan Indiana Cities for 1950 . . . . .	67
18	Cross-Sectional Samples of Non-Metropolitan Indiana Cities for 1960 . . . . .	68
19	Cross-Sectional Sample of Non-Metropolitan Indiana Cities for 1970 . . . . .	68
20	Vertical Displacement of Sample Cities in Reference to a Horizontal Line Thru the Center of Indianapolis . . .	70
21	Assumed Highway Travel Speeds - 1966. . . . .	76
22	Estimated Rural Highway Travel Speeds - 1960 and 1970 .	78
23	Estimation of Average Travel Speeds for Rural Highways in Indiana - 1950 . . . . .	78
24	Assumed Urban Highway Speed for Coding Inter-City Highway Network . . . . .	80
25	Size of Coded Regional Highway Network - 1940 to 1970 .	82
26	Interaction Potential for 1940 Sample Cities. . . . .	85
27	Interaction Potential for 1950 Sample Cities. . . . .	86
28	Interaction Potential for 1960 Sample Cities. . . . .	87
29	Interaction Potential for 1970 Sample Cities. . . . .	88
30	Classification Used in Land Use Measurement . . . . .	92
31	Comparison of Statistically Point Sampling with Inventory Method of Land Use Measurement. . . . .	99
32	Testing Repeatability of Land Use Determination by Statistical Point Sampling of Aerial Photographs. . . .	101
33	Land Use Measurements for 1940 Sample Cities in Acres .	103



## LIST OF TABLES (Continued)

<u>Table</u>		<u>Page</u>
34	Land Use Measurements for 1950 Sample Cities in Acres .	104
35	Land Use Measurements for 1960 Sample Cities in Acres .	105
36	Land Use Measurements for 1970 Sample Cities in Acres .	106
37	Census Data for 1940 Sample Cities. . . . .	108
38	Census Data for 1950 Sample Cities. . . . .	109
39	Census Data for 1960 Sample Cities. . . . .	110
40	Census Data for 1970 Sample Cities. . . . .	111
41	1940 Path Model with Unnormalized Variable and Regression Coefficients . . . . .	116
42	1950 Path Model with Unnormalized Variables and Regression Coefficients . . . . .	121
43	1960 Path Model with Unnormalized Variables and Regression Coefficients . . . . .	126
44	1970 Path Model with Unnormalized Variables and Regression Coefficients . . . . .	131
45	Testing the 1940 Path Model for Predicting Land Use Variables . . . . .	139
46	Testing the 1950 Path Model for Predicting Land Use Variables . . . . .	140
47	Testing the 1960 Path Model for Predicting Land Use Variables . . . . .	141
48	Testing the 1970 Path Model for Predicting Land Use Variables . . . . .	142
49	Independent Variables Used to Explain Variation of Land Uses in Individual Models. . . . .	145
50	Number of Possible Paths Thru Which an Independent Variable May Influence a Dependent Variable in an Eleven Variable Path Model. . . . .	151
51	Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1940. .	153



## LIST OF TABLES (Continued)

<u>Table</u>		<u>Page</u>
52	Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1950. .	154
53	Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1960. .	155
54	Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1970. .	156
55	Summary of Results Obtained from Calibration of 1940, 1950, 1960 and 1970 Path Models . . . . .	166



## LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	Time Path for an Impact Variable - Alternative 1. . . .	14
2	Time Path for an Impact Variable - Alternative 2. . . .	15
3	Fourfold Map of Ends and Means in Applied Statistics .	23
4a	Path Diagram Example Showing "Path Regression Coefficients" . . . . .	27
4b	Path Diagram Example Showing "Path Coefficients". . . .	27
5	Diagrammatic Presentation of Transportation Supply- Demand Theory . . . . .	32
6a	Transportation-City Development Path Model with Unnormalized Variables and Path Regression Coefficients. . . . .	36
6b	Transportation-City Development Path Model with Normalized Variables and Path Coefficients. . . . .	36
7a	Path Model with Unnormalized Variables and Regression Coefficients . . . . .	37
7b	Path Model with Normalized Variables and Path Coefficients. . . . .	37
8	Sampling Scheme for Calibration of Transportation- City Development Path Model . . . . .	42
9	Stratification of Indiana Cities Across Time, Geographic Location, and Population Size. . . . .	44
10	Division of State of Indiana into Three Zones . . . . .	46
11	Counties Excluded (Containing Cities Designated as Metropolitan Areas) from 1940 Data. . . . .	48
12	Counties Excluded (Standard Metropolitan Areas) from 1950 Data . . . . .	49





## LIST OF FIGURES (Continued)

<u>Figure</u>		<u>Page</u>
13	Counties Excluded (Standard Metropolitan Statistical Areas) from 1960 Data . . . . .	50
14	Counties Excluded (Standard Metropolitan Statistical Areas) from 1970 Data . . . . .	51
15	Coded Regional Highway Network Used for Calculation of Interaction Potential - 1940 . . . . .	74
16	Rural Travel Speed 1960-1970. . . . .	77
17	Dot Patterns Designed for Statistical Sampling. . . . .	96
18	Path Coefficients of Calibrated 1940 Transportation City Development Model. . . . .	117
19	Path Coefficients of Calibrated 1950 Transportation City Development Model. . . . .	122
20	Path Coefficients of Calibrated 1960 Transportation City Development Model. . . . .	127
21	Path Coefficients of Calibrated 1970 Transportation City Development Model. . . . .	132



## ABSTRACT

The purpose of this research was to develop a causal quantitative relationship between transportation development and city development.

A generalized eleven variable transportation city development path model was hypothesized, and calibrated separately with 1940, 1950, 1960 and 1970 collected data. Cities sampled were representative of all non-metropolitan Indiana cities with population between 2500 and 50,000.

The individual calibrated models sufficiently explained the variations of land uses within the city, but the four calibrated models were different from each other and could not be combined into one model.

The diagram of calibrated path models and path coefficients were used to calculate total effect coefficients for effect of each significant dependent variable on each independent variable while controlling for extraneous causes.



## CHAPTER ONE

### INTRODUCTION

A new or improved transportation facility influences land development. The resultant land development generates additional trips, which in turn affect the operational characteristics of the transportation facility. The transportation facility and the resultant land development in turn affect the economic development and environmental condition of the region and the community (1)\*. Estimation of probable land development associated with a transportation facility is essential to proper evaluation of a given transportation alternative. Unfortunately, it is difficult to predict probable land development associated with a planned transportation facility. As a result, proper consideration has not been given to this criterion in the transportation system evaluation process. Generally, the resultant land development has been obtained from "intuitively derived generalities" (2).

The exact relationship of transportation and land use is not clear. For example, Kingham documents qualitative evidence that roads do control land use, as well as evidence that roads do not control land use (3). Putman, et al, confirm existence of qualitative evidence of an interrelationship between highway development and land development, but finds no study to show a quantitative causal relationship between

---

\*Numbers refer to items in the Bibliography.



land use and transportation. They state "... , there does not seem to be any prospect of obtaining such evidence since, in the strictest sense this would require, the performance of a controlled experiment on the real world" (4).

It was determined, from studies of the available literature, that the attempt to obtain a quantitative causal relationship between transportation and land use required an approach different from previously performed transportation-land use studies. Previous studies generally tried to measure the effect of change on the land use which came about as a result of change in the transportation system. As will be more fully explained in Chapter 2, this method, normally called a "before-and-after study", belongs to the family of experiments that requires control. Basic to a controlled experiment is the ability to replicate the experimental units. It is not possible to replicate an urban environment and the subsequent transportation change. Therefore, studies using variations of "before and after" studies have not been successful in explaining the interrelationship of transportation and land use.

Fortunately, replication of experimental units (urban areas) is not necessary for learning about the generalized relationship between transportation and land use. "Quasi-Experimental Design" allows learning from observation of experimental units without having replication of experimental units. "Design of Experiment" is that part of statistical science which allows, for example, agricultural scientists to compare the quality of two seeds by keeping all other conditions constant (by replicating experimental units). "Quasi-design of





Experiment" uses the technique of "Design of Experiment" for occasions when replications of experimental data are not possible.

### Purpose

The purpose of this research is to develop a generalized quantitative causal relationship between transportation development and urban development. This quantitative relationship is useful in the evaluation phase of the transportation process, as well as in transportation impact studies.

### Scope

The inference space of this study was limited to cities in Indiana with a population greater than 2500 and less than 50,000, which were not located within standard metropolitan statistical areas (SMSA). The data used in this study was collected from four study periods of 1940, 1950, 1960 and 1970.

The lower limit of 2500 was chosen because certain census data were not available for smaller cities. An upper limit of 50,000 was chosen because in this study, land uses were measured from aerial photographs in terms of acreage of land use for a given purpose. It was felt that in larger towns, due to density of structural pattern, square footage of different land uses had to be measured instead of land uses in acres. Measurement of floor space was costlier than measurement in acres, and historical measurements of floor space was not readily available (although possible from aerial photographs).

The inference space was also limited to cities in non-SMSA counties. This limitation was placed in order to avoid dealing with



the problem of the experimental city being adjacent to a larger city, or in an integrated and built-up area.

The model(s) developed are aggregate models in the sense that the city itself is the experimental unit. The models developed measure the change in a particular land use for the city, as a whole, as a result of change in the transportation system. The developed model is not sensitive for land use changes in individual traffic zones.

#### References

1. Manheim, Marvin L., et al, "Transportation Decision-Making - A Guide to Social and Environmental Considerations", National Cooperative Highway Research Program, Report 156, 1975, p. 156.
2. Frye, Frederick, "Alternative Multimodal Passenger Transportation Systems: Comparative Economic Analysis", National Cooperative Highway Research Program, Report 146, 1973, p. 12.
3. Kingham, Ian R., "Suburban Highways and Roads as Instruments of Land Use Change", Transportation Research Board, Record 565, 1976, p. 54.
4. Putman, Stephen H., The Interrelationship of Transportation Development and Land Development, prepared for Federal Highway Administration, 1973, Vol. 1, p. 2.



## CHAPTER TWO

### PREVIOUS INVESTIGATION

Transportation impact studies number in the thousands (1). A few of these reports were studied in the early part of the research (2 to 12). These reports provided an understanding of impact studies which would have been hard to obtain from studies of abstracts and literature reviews alone. The sheer number of transportation impact studies, and the fact that many of the studies are repetitive, necessitated extensive use of available state-of-the-art reviews of transportation impact studies (1, 13 to 20).

#### Types of Transportation Impact Studies

Transportation impact studies were categorized as to type of transportation facilities studied to determine methodologies used and results obtained. Putman, et al, divide transportation impact studies into the following five general categories (21):

1. Impact of highway bypasses.
2. Impact of new highway construction on a narrow corridor.
3. Land development around freeway interchanges.
4. Impact of major highway construction upon a region.
5. Relationship of transportation development and land use development.



### By-Pass Studies

Land use changes were reported for many by-passes. There were some excellent early bypass studies. Unfortunately, bypass studies become repetitious, and generalized knowledge useful in prediction of land use consequences of a particular bypass was not available.

In 1949-50 a 7.1 mile long, free access, two-lane bypass was built near Kokomo, Indiana. Kokomo (1950 population 38,600) is a "typical midwestern city that has grown and prospered in the last one hundred years". A 1951 study of the Kokomo bypass shows the land use maps for 1938 and 1948 obtained from aerial photographs, and a land use map for 1951 obtained from a field study. These land use maps show that prior to 1948 the growth of the city was toward the south and southwest. In 1951, the growth continued toward the south and southwest, but there was also growth along the bypass (23). A re-study of the same bypass in 1965 showed the shift in direction of city growth from the south and southwest to the east where the bypass was located (24).

Vargha (25) in a study of 218 small communities discussed their growth in relation to highway bypasses and natural barriers and concluded bypasses with limited access act as natural barriers. When one side of the community was contained by a natural barrier and a limited access bypass was contemplated for the other side, he recommended enough land be left between existing community and the bypass for the community to grow to a population size of 1200 to 2000. If this margin of land for potential growth was not included, the community remains so small that it acts as one unit and cannot cross the barrier of an access-limited bypass.





Skorpa, et al, summarize the results of land use changes from before-and-after studies of four highway bypasses in the state of Texas. They conclude that the main effect of these bypasses was to change agricultural land into land held for future use. Also they state that the effect of land use changes was basically limited to a "very narrow strip along the facility" (26).

#### Impact of Radial Highways on a Narrow Corridor

Generally, the impact of radial highways was measured in terms of increase in land value for property close to the roadway. Also, in urban areas, the land use generally was changed to commercial land use (27). References 12 to 15 describe several of these studies. One typical example of this type of impact study was "a before-and-after" study of effects on land values and land use of 5.4 miles of the Dallas central expressway.

The Dallas study, in part, collected land use and land values for parcels close to the expressway, from three different "bands". The "A band" included properties abutting the expressway's frontage road. The "B" and "C" bands which were two blocks wide parallel the "A" band. There was also a control area. Land values in the three bands were compared with land values in the control areas. The conclusion reached was that land values adjacent to the expressways increased more than land values in the control areas (28).

#### Interchange Studies

Skorpa, et al, summarized the results of studies of land development adjoining 66 interchanges in Michigan, 36 interchanges in



Pennsylvania, 221 interchanges in North Carolina and 102 interchanges in Indiana. Their conclusion was that "despite many similarities in land development of different interchanges locations, too little is known about the underlying factors to be able to predict future land use at a specific interchange" (29).

#### Impact of Major Highway Construction on a Region

Another type of highway impact study was the attempt to find the effects of major highway improvement on a region. Objectives of these studies were similar to that of the current project. Unfortunately, choice of improper methodology made it difficult to obtain generalized knowledge applicable to the planning of future transportation projects. An early example of this type of study was the economic impact study of Massachusetts Route 128 (30). Two additional examples were studies of impact of freeways in the states of North Carolina (31) and Connecticut (32). The North Carolina study was one of the more ambitious studies of this group and attempted to develop predictive land development equations.

The North Carolina study attempted to measure the impact of rural, suburban and urban freeways in the state of North Carolina separately. Average traffic volumes and land use for 310 miles of rural highways with 105 interchanges as well as 171 miles of suburban freeways with 76 interchanges was inventoried. The inventories and subsequent analysis lead to "several general observations relative to the highway facilities in the rural and suburban areas" (33).



The most comprehensive analysis in the North Carolina study was that of land development characteristics of the Raleigh Beltway (34). Individual multiple regression equations using "before and after" data were developed to predict the following: (1) the number of housing and apartment units, (2) the value of industrial developments, (3) commercial developments, and (4) office and institutional developments for Raleigh, North Carolina. The value of the coefficient of determination ( $R^2$ ) for the developed equations were generally satisfactory ( $R^2 = 0.7$  to  $0.9$ ).

The models generally correlate the positive impact of the Raleigh beltway with adjoining land use, but do not include consideration of any possible negative impact of the beltway on other parts of the Raleigh community. The study shows that some land values close to the beltway have increased from \$3,000 per acre to \$50,000 per acre. The study also states "land values have gone down significantly in the core area, and surrounding the core area, the developments are of very low rent type facilities with a high vacancy ratio" (35), but this phenomenon was not considered in the model development.

#### Interrelationship of Transportation and Land Development

There were only a few studies which attempted to find the underlying interrelationship between transportation and land development. One such attempt was that of Bascom, et al, who after a review of the literature, hypothesized an approximate causal model compromised of a set of five simultaneous regression equations to estimate "secondary effect" of investment in highways and sewers in urban development. Unfortunately they did not calibrate their hypothesized model but



reduced the formulated five equation model to a single equation model for operational purposes (37). The data set, borrowed from calibration of the "Empiric land use model" in Boston, Denver, Minneapolis-St. Paul, were then used to calibrate the single equation model. The calibration process resulted in a low level of explanation ( $R^2 = 0.17$  to  $0.49$ ). The general conclusion was that the effect of highways on single family residential construction was more significant than originally assumed and the effect of sewer investment was very important. However, because of unstability of model parameters, the model interpretations were questionable (37).

#### Summary of Previous Transportation Impact Studies and Their Methodology

A review of literature of transportation impacts on land use has not produced any generalized causal, quantitative tools for estimating land use consequences of a transportation system change on a community. It is believed the problem lies in the fact that the same "before and after" methodology is used for one transportation impact study after another. Another basic problem has been the general implicit assumption that the effect of highways on land use is limited to a narrow corridor. As a result of this last assumption, a general increase in land values, adjacent to the facility, has been measured as a positive impact. There has been no attempt to measure the net highway effect by assessing possible negative effects of the facility elsewhere, particularly on the central business district. The difficulty with before and after methodology and lack of comprehensiveness in impact studies will be





discussed in the following section and a framework for an improved alternative will be recommended.

### Analysis of Transportation Impact Studies Using "Before-and-After" Methodology

The method of "before-and-after" analysis was used to measure the effect of a new or improved highway on land values, land use and other socio-economic variables. The variables of interest were usually measured two to five years prior to facility improvement and then again two to five years after the change. It was recognized in most studies that there were other influences besides transportation variables which affect the socio-economic variables of interest. Therefore, a control area was chosen to isolate the effect of the transportation change from the effects of these exogenous variables on the study area. In absence of a suitable control area, multiple regression analysis was sometimes used to relate one dependent variable to several independent variables (38).

Generally, an important goal of transportation impact studies was to determine the cause and effect relationship between transportation development and land development so as to assist decision makers in future planning of transportation facilities. "In truth, policy making bodies cannot have much more ability to decide between alternative transportation investments after reading these studies than before" (39). The problem was due to "completely inadequate theoretical foundation and methodological approaches" (40). Intuitively, "before-and-after" techniques may seem appropriate for measurement of a transportation impact, but their use was the one essential methodological



problem with transportation impact studies. Also, the associated choice of control area as an attempt to isolate the impact of transportation change from that of other exogeneous influences was misleading. The urban area subjected to transportation improvement was systematically subjected to the particular transportation change because it was in some ways different than other areas, including the control area.

The use of "before-and-after" methodology to measure the impact of a new or improved transportation facility, at least implicitly, assumes that transportation change affords a unique opportunity for performing an "experiment" to determine the effects of the transportation change upon the community. The above assumption was in the tradition of "design of experiment". Without explicitly stating it a "before-and-after" study assumes that the transportation improvement was a "treatment", and that the affected area, be it a narrow strip of land, the whole urban area, or an entire region, was an "experimental unit". Essential to the methodology of "design of experiment" was the ability to replicate the "experimental units".

In the physical sciences "control" of experimentation was readily available. In such sciences as agriculture and biology as well as many aspects of engineering and manufacturing, "control" of experiment was obtained through random assignment of treatments into replicated experimental units. Transportation impact studies present a greater challenge than the above studies because it is not possible to replicate experimental units. An urban environment is very heterogeneous and it is not possible to divide urban areas into similar experimental units. Also, treatment is not assigned randomly to the "experimental



units", "for the treatment, the transportation change is always designed to relate systematically to some aspects of distribution of activities and residents" (41). In the above discussion "before-and-after" methodology is criticized on a theoretical basis. It is useful to envision the potential of learning from "before-and-after" methodologies, for the moment assuming that the theoretical difficulties are overcome. Figures 1 and 2 show two possible scenarios for time variations of one socio-economic variable with and without transportation improvement.

The scenario represented in Figure 1 yields more information in conjunction with "before and after" study than many other possible scenarios, including the one shown in Figure 2. In these Figures, curve X represents the time value of one variable as influenced by the transportation improvement, and curve Y represents the corresponding value of the same variable without the transportation improvement. Point C represents the time at which the transportation change took place and points B and A represent times of "before" and "after" data collection.

From measurement of the variable of interest (for example, average land value) before (B) and after (A) the study area is subjected to the particular transportation change, the quantity  $X_A - Y_B$  is obtained. In the case of the scenario presented in Figure 1, the quantity  $X_A - Y_B$  is positive, but unfortunately this increase is not attributed to the change in transportation, because the transportation variable is hopelessly "confounded" with other variables. Variables, besides transportation, which may affect urban growth are numerous. For



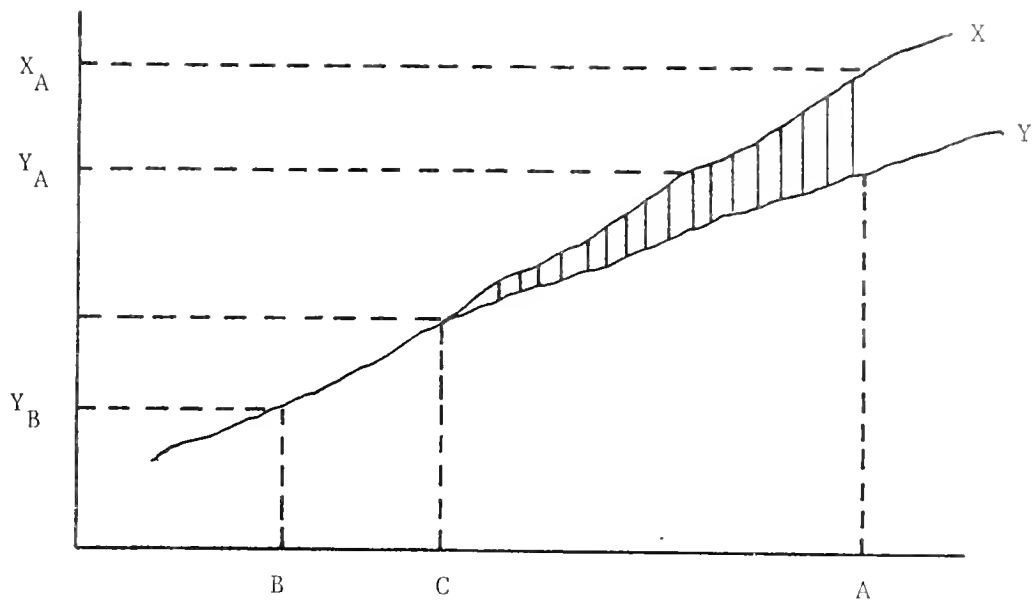


Figure 1. Time Path for an Impact Variable - Alternative 1

Source: Charles River Associates, Inc., Measurement of the Effects of Transportation Changes, Prepared for Department of Transportation, Sept. 1972.





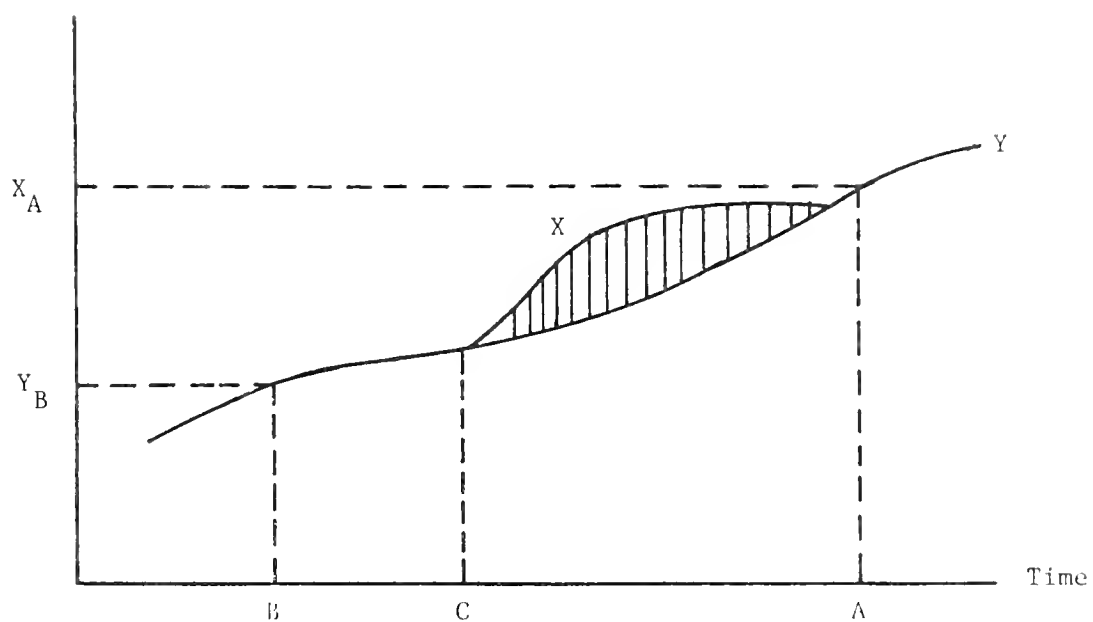


Figure 2. Time Path for an Impact Variable- Alternative 2

Source: Charles River Associates, Inc., Measurement of the Effects of Transportation Changes, Prepared for Department of Transportation, Sept. 1972.



example, economic condition of the region, population, educational opportunities and location of the urban area probably have significant effects on the measured,  $X_A - Y_B$ , quantity. The effect of transportation change on the measured variable could have actually been negative, and the positive value measured may be due to exogenous variables.

To overcome the difficulty caused by "confounding" of transportation impact with other variables, many researchers have chosen a "control area" in conjunction with "before-and-after" design to separate the impact of transportation improvement from the impact of other variables. The "control area" is chosen for its similarity with the "study area" in all aspects except for the fact that transportation improvement takes place in the "study area", not in the "control area". It is then assumed that exogenous variables affect both "control area" and "study area" in an equal manner, and any measured differences between the value of the variables of interest in the "study area" and the "control area" are due to transportation change.

Assuming for now, that the proper "control area" as described in the above paragraph is selected. Then the magnitude of change in the variable of interest, without transportation change is measured. In Figure 1, this value is represented by quantity  $Y_A - Y_B$  which allows adjustment of the previously measured quantity,  $X_A - Y_B$  to obtain the quantity  $X_A - Y_A$  which is the net change in the variable at the time A. This last quantity shows whether the transportation impact is negative or positive, but it is not a measure of magnitude of the transportation impact, because transportation impact is represented by the dashed area in Figure 1.



The scenario presented in Figure 1 is ideal for yielding information in conjunction with "before-and-after" methodology. For example, the situation presented in Figure 2 does not allow measurement of the impact of transportation improvement at all, because the impact of the transportation change was temporary and the interval between transportation change (C) and the observation (A) was too long (41). The above analysis shows that the maximum amount that is learned from traditional "before-and-after" studies with proper control areas is whether the transportation impact is positive or negative.

Urban areas are heterogenous. The fact that the urban area is subjected to the transportation change, and the control area is not, is proof enough that the two urban areas are different from each other. Two urban areas at one point in time seem to be similar in some aspects but most probably these cities take diverse growth patterns (42).

#### References

1. Charles River Associates, Inc., Measurement of the Effects of Transportation Changes, prepared for DOT, Sept. 1972, p. 2-1.
2. Michael, Harold L., "An Economic Evaluation of the Kokomo Indiana By-Pass", M.S. Thesis, Purdue University, August 1951.
3. Branham, A. K., et al, "Economic Evaluation of Two Indiana Bypasses", Highway Research Board, Bulletin 67, 1953, pp. 1-14.
4. Putman, Stephen H., et al, The Interrelationship of Transportation Development and Land Development, prepared for Federal Highway Administration, 1973, Vol. I, p. 2.
5. Kelly, J. F., "Tulare Bypass", Reprint from California Highway and Public Work, May-June 1956.
6. Harrison, Joseph W., "Economic Studies Progress Report No. 2 Study of the Economic Effects of the U.S. Route 11 ByPass at Lexington, Virginia on Business Volume and Composition", Virginia Council of Highway Investigation and Research, Charlottesville, Virginia, 1958.



7. Horwood, Edgar M. and Boyee, Ronal R., "Measurement of Central Business District Change and Urban Highway Impact", Highway Economic Studies, University of Washington, 1959.
8. Fletcher, Joseph A., "Early Effects of Rural Interstate Highways", M.S. Thesis, Purdue University, June 1961.
9. Lohr, Alan F., "Early Impact of a Highway Improvement on an Urban Area", M.S. Thesis, Purdue University, 1962.
10. Texas Transportation Institute, "Economic Impact Study - Waxahachie, Texas", Bulletin No. 35, March 1966.
11. Fabbroni, Lawrence P., "Land Use Development at Interstate Interchanges in Indiana", Joint Highway Research Project, No. 9, May 1973.
12. Brown, F. A., and Michael, H. L., "The Impact on Land Value of a Major Highway Interchange Near a Metropolitan Area", Joint Highway Research Project, No. 34, December 1973.
13. Bureau of Public Roads, Highways and Economic and Social Changes, 1964.
14. Federal Highway Administration, Economic and Social Effect of Highways, 1972.
15. Federal Highway Administration, Economic and Social Effect of Highways, 1974.
16. Federal Highway Administration, Social and Economic Effects of Highways, 1976.
17. Skorpa, Liverd, et al, Transportation Impact Studies: A Review with Emphasis on Rural Areas, Texas University at Austin, NTIS PB-239463, 1974.
18. Putman, Stephen H., et al, Vol. I and II.
19. Bascom, S. E., et al, Secondary Impacts of Transportation and Wastewater Investments: Review and Bibliography, U.S. Environmental Protection Agency, Report No. EPA 600/5-75-002, January 1975.
20. Isibor, E. I., "Modelling the Impact of Highway Improvements on the Value of Adjacent Land Parcels", Ph.D. Thesis, Purdue University, January 1970.
21. Putman, Stephen H., et al, pp. 6-8.
22. Michael, Harold L., p. 46.
23. Ibid, p. 129.





24. Evans, Eugene Glem, "Impact of Kokomo By-Pass from 1950 to 1964", M.S. Thesis, Purdue University, August 1965, p. 97.
25. Vargha, Louis A., "Highway Bypasses, Natural Barriers and Growth in Michigan", Highway Research Board, Bulletin, 250, 1960, pp. 29-36.
26. Skorpa, Livard, et al, "Transportation Impact Studies", pp. 19-20.
27. Putman, Stephen H., et al, Vol. I, p. 7.
28. Adkins, William G., "Effects of the Dallas Central Expressway on Land Values and Land Use", Texas Transportation Institute, Sept. 1957.
29. Skorpa, Livard, et al, pp. 9-15.
30. Bone, A. S. and Wohl, M., "Massachusetts Route 128 Impact Study", Highway Research Board, Bulletin 227, 1959.
31. Babcock, W. F., "An Analysis of the Impacts of Freeways on Urban Land Developments in North Carolina (Including Guidelines for Highway Planning and Design)", North Carolina State University, June 1974.
32. Hammerschlag, Dieter, et al, "The Interstate Highway System and Urban Structure: A Force for Change in Rhode Island", University of Rhode Island, July 1976.
33. Babcock, W. F., pp. 25-29.
34. Ibid, p. 37, and 109-115.
35. Ibid.
36. Bascom, S. E., et al, Secondary Impacts of Transportation and Wastewater Investments: Research Results; U.S. Environmental Protection Agency, Report No. EPA 600/5-75-013, pp. 15-17.
37. Ibid, pp. 18-61.
38. Skorpa, Livard, et al, pp. 5-7.
39. Charles River Associates, Inc. pp. 28-29.
40. Ibid.
41. Ibid, Chapter 4, p. 25.
42. Ibid, Chapter 2, p. 23.



### CHAPTER THREE

#### A FRAMEWORK FOR DEVELOPING THE INTERRELATIONSHIP OF TRANSPORTATION DEVELOPMENT AND CITY DEVELOPMENT

The difficulty discussed in previous transportation studies arose due to lack of "control" over the experimental conditions. It is possible to learn about the impact of transportation without the performance of experiments. Whereas in agriculture science, replication of experiment is necessary, in transportation impact analysis:

' "We must learn principally from observation without experimentation. We are closer to the situation of geophysicists postulating and testing the theory of continental drift than to Fisher studying, designing, and interpreting agricultural experiments" (1).

"Quasi-experimental" research design allows learning from observations where formal scientific "design of experiment" is not possible. "Quasi-experimental" design uses techniques of "design of experiment" to allow certain statistical control of the phenomenon under study where deliberate randomization of "experimental units" and treatments are not possible. Campbell and Stanley developed this design for use in the field of education (2), and Charles River Associates, Inc. extended Campbell and Stanley's finding to transportation impact problems.

Randomization schemes in "design of experiment" allow tests of significance for a given treatment for the chosen "inference space". The "quasi-experimental" designs are sampling schemes which allow choice



of samples representing the "inference space". The variable of interest is measured from each chosen sample, which in turn allows general statements in regard to significance of the measured variable over the "inference space".

In this chapter "quasi-experimental" design was used to develop an alternative framework to "before and after" design for impact of transportation facilities. The proposed framework does not measure the change brought about as a result of a change in transportation, but attempts to develop a model to explain the relationship of transportation development and city development. The ability to explain this relationship can, in turn, be used to choose between transportation alternative by predicting the consequence of implementation of a given alternative.

The proposed framework for establishing a relationship between transportation development and city development is conceptually simple, and applicable to the study of a wide range of impact problems.

The following four steps are sequential:

1. Development of structural transportation model(s) which specify relationships among a set of variables based on explicit hypothesis.
2. Based on the principles of "quasi-experimental" design, samples representing the desired "inference space" are obtained and the values of the variables in the model are measured.
3. Measured variables are used to calibrate the structural model.
4. The model is tested.



The framework for development of a model which explains the relationship between transportation development and city development is detailed in the following sections.

### Use of Path-Analysis for Verification of Relationship Between Transportation Development and City Development

The task is to specify a structural model which explains the interrelationship of transportation development and city development. Wold (3) describes requirements for "explanation" based on observed "quasi-experimental" (or as Wold calls it, "non-experimental") data. For further clarification, it is useful to study the position of "explanation" from "observed" data in the realm of the science of applied statistics. In Wold's terminology, there are two "means" and two "ends" in applied statistics resulting in "Four-Fold Map of Ends and Means", shown in Figure 3. One of the two "means" is statistical data, obtained from "experimental" designs, and the other is statistical observations, obtained from "non-experimental" designs. The two "ends", or purposes of statistical inference, are "description" and "explanation" (4).

Clarification of the distinction between descriptive statistics and explanatory statistics is useful. Descriptive statistics, which are limited to the northern sector of Figure 3, deal with the questions of "How"? or "How much"?, while explanatory statistics limited to the southern half "treat the question 'Why'? and ... the answer involves an element of causal inference" (5). Descriptive statistics is sufficient to maintain control over current operations, but for improvement of the operations, explanatory statistics are required. A good





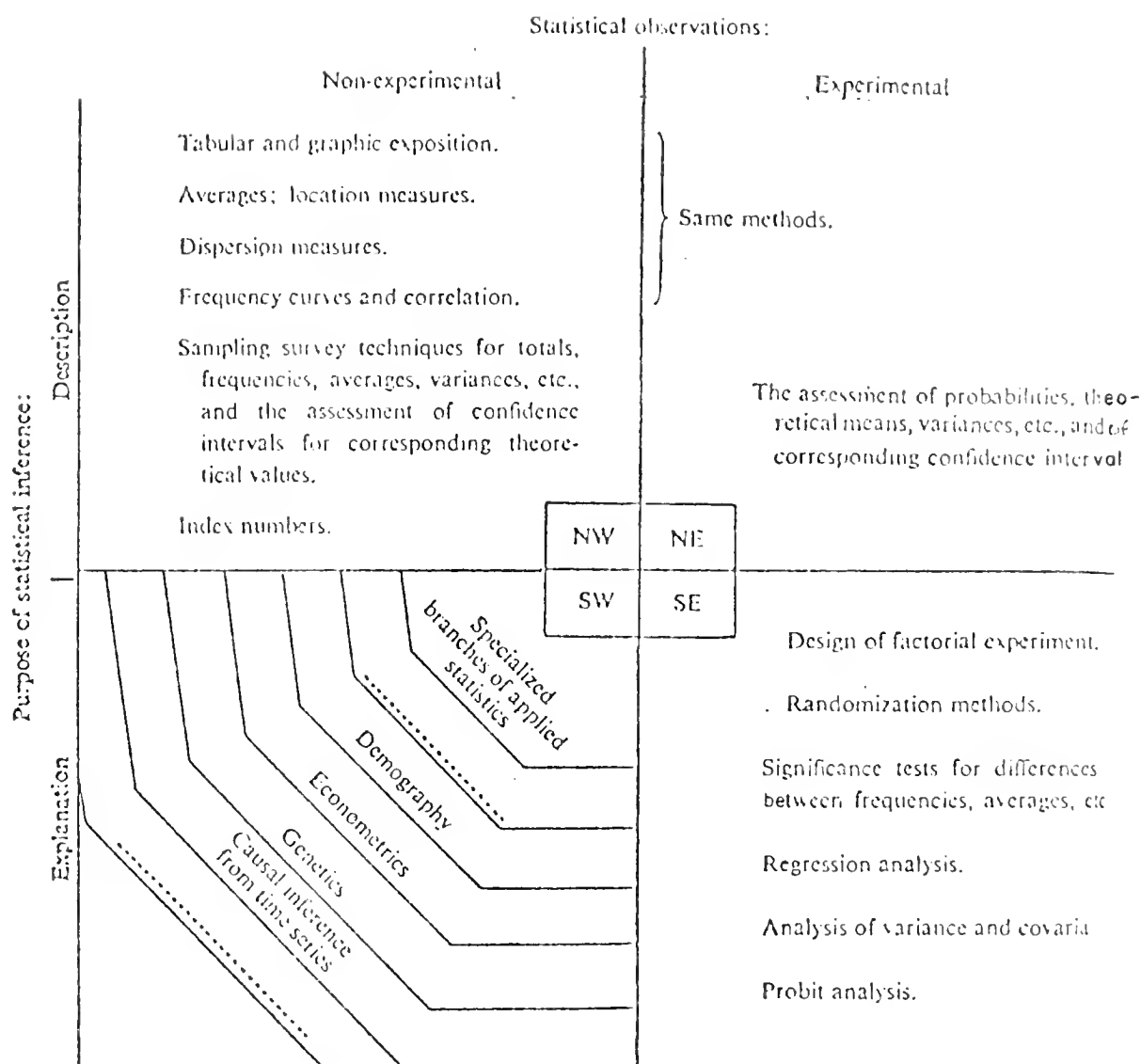


Figure 3. Four Fold Map of Ends and Means in Applied Statistics

Source: Wold, Herman "Causal Inference from Observational Data: A Review of Ends and Means", Journal of Royal Statistical Society, Sec. A., 1956, pp. 28-51.



interpretation of distinction is credited to Dr. R. Porter "... description is employed as an aid in the human 'adjustment' to given condition, while explanation is a vehicle for 'ascendancy' over the environment" (6).

Prior to the 20th century, statistics was largely limited to a description based on "non-experimental" data, while the classic and modern statistics developed in this century was largely interested in description and explanation by use of "experimental" data (NE and SE sectors) (7). According to Wold:

"The epoch-making contributions of R. A. Fisher brought a powerful expansion in the direction of experimental methods, with emphasis in the SE sector. The exact and efficient methods developed by Fisher and his followers have led to great triumphs in biological sciences, medicine, engineering and other fields where research can be based on controlled experiments..." (8).

The progress in the SW sector is slow because "explanation" on the basis of "quasi-experimental data" is more difficult than "explanation" based on "experimental data". Unlike the progress in the SE section, brought by Fisher's formalized and generalized statements applicable to all fields, the progress in the SW section has to come separately for each field. With "non-experimental" data "the statistical analysis has to be closely coordinated with the subject matter, both in specifying the causal hypothesis, and testing them against other sources of knowledge" (9). The specification of the causal model is based on subject matter theory to "test or demonstrate the causal assumptions" (10).

The general structural model chosen in this research project to explain the relationship of transportation and city development was the



"path model". Path analysis was first developed by geneticist Wright in 1918, and popularized in sociology by Duncan in 1966 (11). Path analysis was a useful tool in the applied statistics where the objective was "explanation" from "quasi-experimental" data (SW sector of Figure 3). The path model, as normally used, is a system of recursive regression equations. There are several advantages in the use of the path model. One advantage of path analysis is that the assumptions made are explicit, which in turn forces the discussion to be internally consistent, so that incompatible assumptions are not incorporated within the model. Another advantage of the path model is that effects of one variable on another variable are divided into direct and indirect effects. This in turn allows a measure of the total effect of one variable on another.

Path analysis, as developed by Wright, is concerned with linear, additive, and causal relationships among a set of variables which are measured by an "interval scale". Although the occasional use of a dummy variable is allowed, "the hypothetical variables (dummy variables) are conceived as continuous variables" (12). The assumption of a causal relationship is the only assumption not required of the usual statistical model. "Causal relationship" in path analysis means that a "weak causal order" exists between the set of variables, and, the relationship among these variables are "causally closed". "Weak causal order" between variables X and Y ( $X \rightarrow Y$ ) means that X can affect Y, but Y cannot affect X. By "causal closure", it is meant that given a bivariate covariation, and weak causal ordering between X and Y, then the observed covariation between X and Y is due (1) to causal dependence of



Y on X, (2) to the mutual dependence of X and Y on other variable(s), or (3) to the combination of the above two causes of variations (13).

The diagrammatic representation of the set of structural equations which is based on the above assumptions is called a "path diagram". A "path diagram" is useful in understanding the assumptions made in the "path model", because there is a one to one relationship between statistical and algebraic properties of the structural equations and the "path diagram" (14). The following four conventions are used for presentation of a "path diagram":

1. Uni-directional straight arrows identify "weak causal ordering" of variables. Straight arrows are drawn from each "determining variable" to each variable dependent on it.
2. The non-causal correlated relationship between two independent variables is shown with a two headed curvilinear arrow.
3. The portion of variation of any dependent variable not explained within the system is represented with a "residual variable". The "residual variables" are also shown with a unidirectional arrow.
4. Beside each unidirectional arrow the value of "path coefficient" ( $P_{ji}$ ) is entered. The first subscript  $j$  denotes the dependent variable and the second subscript,  $i$ , represents the independent variable.

A simple example of the "path model" of Figure 4a is offered for visualization of the above conventions, and to demonstrate the technique of "path analysis". The "path estimation equations" are derived from





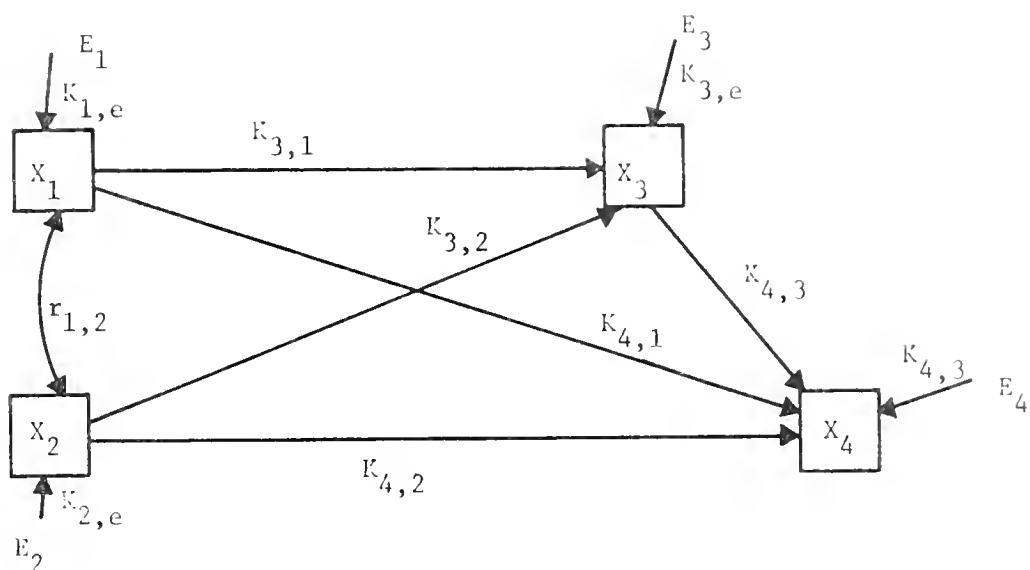


Figure 4a. Example of Path Model with Unnormalized Variables and Path Regression Coefficients

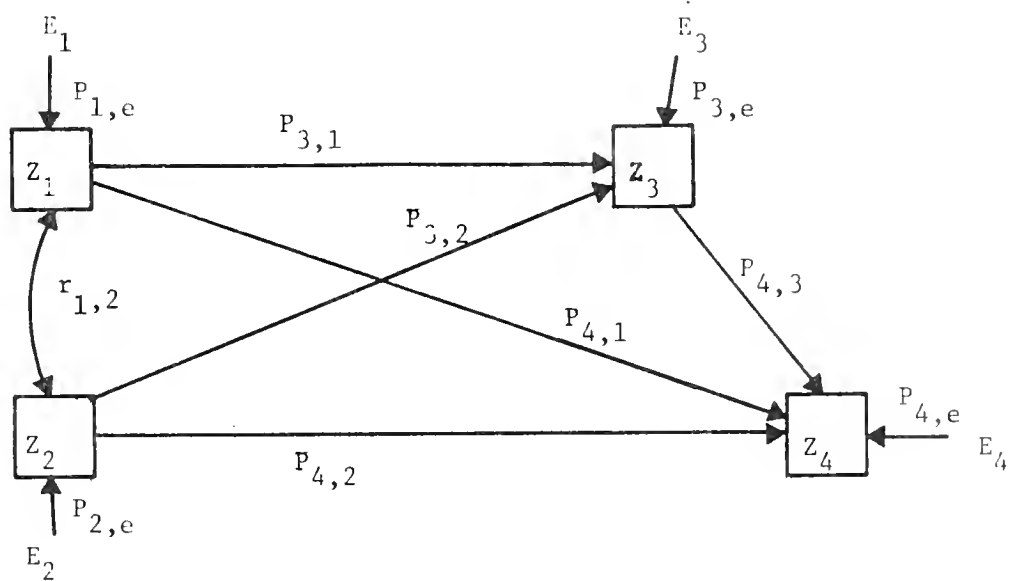


Figure 4b. Example of Path Model with Normalized Variables and Path Coefficients



the structural equations of the "path model" by use of the "path analysis" technique (15). The example also shows how the total effect of an independent variable on a dependent variable is obtained. The general method used in conjunction with solution of the four variable model of Figure 4a is directly applicable to the much larger model, which is used in this research. But the solution of this small model more conveniently demonstrates the technique of "path analysis".

Given the four variables  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_4$ , and the "path model" of Figure 4a, the structural equations are:

$$X_3 = K_{3,1} X_1 + K_{3,2} X_2 + K_{3,e} E_3 \quad \text{eq. 1a}$$

$$X_4 = K_{4,1} X_1 + K_{4,2} X_2 + K_{4,3} X_3 + K_{4,e} E_4 \quad \text{eq. 1b}$$

The K coefficients are known as path regression coefficients, and are obtained by the ordinary method of regression equation. For example, the value of  $K_{3,1}$  is the same as the regression coefficient obtained by regressing  $X_3$  on  $X_1$ , when  $X_2$  is already in the equation. This last statement in regression terminology is written as  $\beta_{31.2}$ .

If the mean and variance of the variable  $X_i$  are  $M_i$  and  $\sigma_i$ , then  $Z_i = (X_i - M_i)/\sigma_i$  represents the variable in the form of its "standard score". "Path coefficient",  $P_{ji}$ , between two standardized variables with subscripts i and j is defined as  $P_{j,i} = K_{j,i}(\sigma_i/\sigma_j)$ . Then by substitution equation 1a and 1b are transformed to equation 2a, and 2b.

$$Z_3 = P_{3,1} Z_1 + P_{3,2} Z_2 + P_{3,e} E_3 \quad \text{eq. 2a}$$

$$Z_4 = P_{4,1} Z_1 + P_{4,2} Z_2 + P_{4,3} Z_3 + P_{4,e} E_4 \quad \text{eq. 2b}$$



The actual values of the "path coefficient" is calculated in the same way as path regression coefficients, but the variables are in standardized form. Figure 4b represents the path model with standardized variables and path coefficients. The value of residuals for variable  $Z_i$  is  $\sqrt{1-R^2}$ ; where  $R^2$  is the value of the "coefficient of determination" obtained when  $Z_i$  is the dependent variable, and all independent causal variables are in the equation.

Path coefficients are statistically tested in accordance with backward elimination method. In case the value for one of the coefficients is negligible the corresponding path is removed from the "path diagram", and the process is repeated until all the remaining path coefficients are significant.

#### Total Effect of an Independent Variable on a Dependent Variable

To obtain the "total effect" of variable  $Z_i$  on variable  $Z_j$  within the "path model", let's determine equations of type:

$$Z_j = C_{ji} Z_i \quad \text{eq. 3}$$

$C_{ji}$  is a constant standing for magnitude of change in  $Z_j$  for a unit change (i.e. a one standard deviation change) in  $Z_i$ , while controlling for extraneous causes. For the path model of Figure 4b:

$$C_{3,1} = P_{3,1} \quad \text{eq. 4a}$$

$$C_{3,2} = P_{3,2} \quad \text{eq. 4b}$$

$$C_{4,3} = P_{4,3} \quad \text{eq. 4c}$$

$$C_{4,2} = P_{4,2} + P_{3,2} P_{4,3} \quad \text{eq. 4d}$$

$$C_{4,1} = P_{4,1} + P_{3,1} P_{4,3} \quad \text{eq. 4e}$$



The value of  $C_{4,2}$  obtained in equation 4d, represents the total effect of variable  $Z_2$  on variable  $Z_4$ . This total effect is a summation of the direct effect of variable  $Z_2$  on variable  $Z_4$  and an indirect effect of variable  $Z_2$  on variable  $Z_4$ , thru variable  $Z_3$ . The interpretation of the term  $C_{4,2}$  is that for one standard deviation change in variable  $X_2$ , variable  $X_4$  is changed  $C_{4,2}$  standard deviation. The concept is very helpful in a more complex model where there are many indirect effects, making intuitive assessment of total effects almost impossible.

The above explanation of "path analysis" is a mechanical description of the subject which is highly dependent on subject-matter assumptions for its usefulness. In the following paragraphs, a "path model" is hypothesized to explain the relationship between transportation and city development. In the following chapters the method of "path analysis" is incorporated in a quasi-experimental design to calibrate the hypothesized model.

#### Path Model Hypothesized to Explain Relationship Between Transportation Development and City Development

Development of a "path model" to explain the relationship between transportation development and city development depends upon subject matter theory and assumptions.

The most essential assumption in the path model developed in this project is that transportation development influences city development. This assumption is based on transportation "supply and demand" theory, which states in the long run the transportation improvement is influenced by major investment decision making.





Transportation "supply and demand" theory which is based on welfare economics is discussed elsewhere (16). A simplified explanation of this theory with help of the diagrammatic presentation of Figure 5 is given here. Curves DD and SS represent the initial demand and supply functions. The initial equilibrium is at Point  $E_1$ ; with a trip volume of  $T_1$  and corresponding trip cost of  $C_1$ . A transportation improvement shifts the supply function to the right, to the  $S_1S_1$  position. The shift of the supply function to the right if the demand for travel remains the same as the pre-improvement period causes a much lower travel cost  $C_2$ , but this is an unstable state and shortly after the improvement drivers divert from other routes to take advantage of the lower travel cost in the improved route until equilibrium  $E_3$  with corresponding trip volume  $T_3$  and trip cost  $C_3$  is reached. In the long run though, due to lower travel cost  $C_3$ , investment decisions are made to relocate or initiate new construction to take advantage of the improved transportation facility. These investment decisions, in turn, cause the demand curve to shift to the right which means a higher travel cost for a given trip volume. Eventually the equilibrium point between a new supply function and a new demand function is reached at point E with the corresponding trip volume T and cost C. The consequence of this theory is that transportation development influences city development.

The path model developed in this research project is based on the eight specific assumptions listed in Table 1. Assumptions one and two partially follow from transportation supply and demand theory. Assumptions three thru eight are not based on any one specific source,



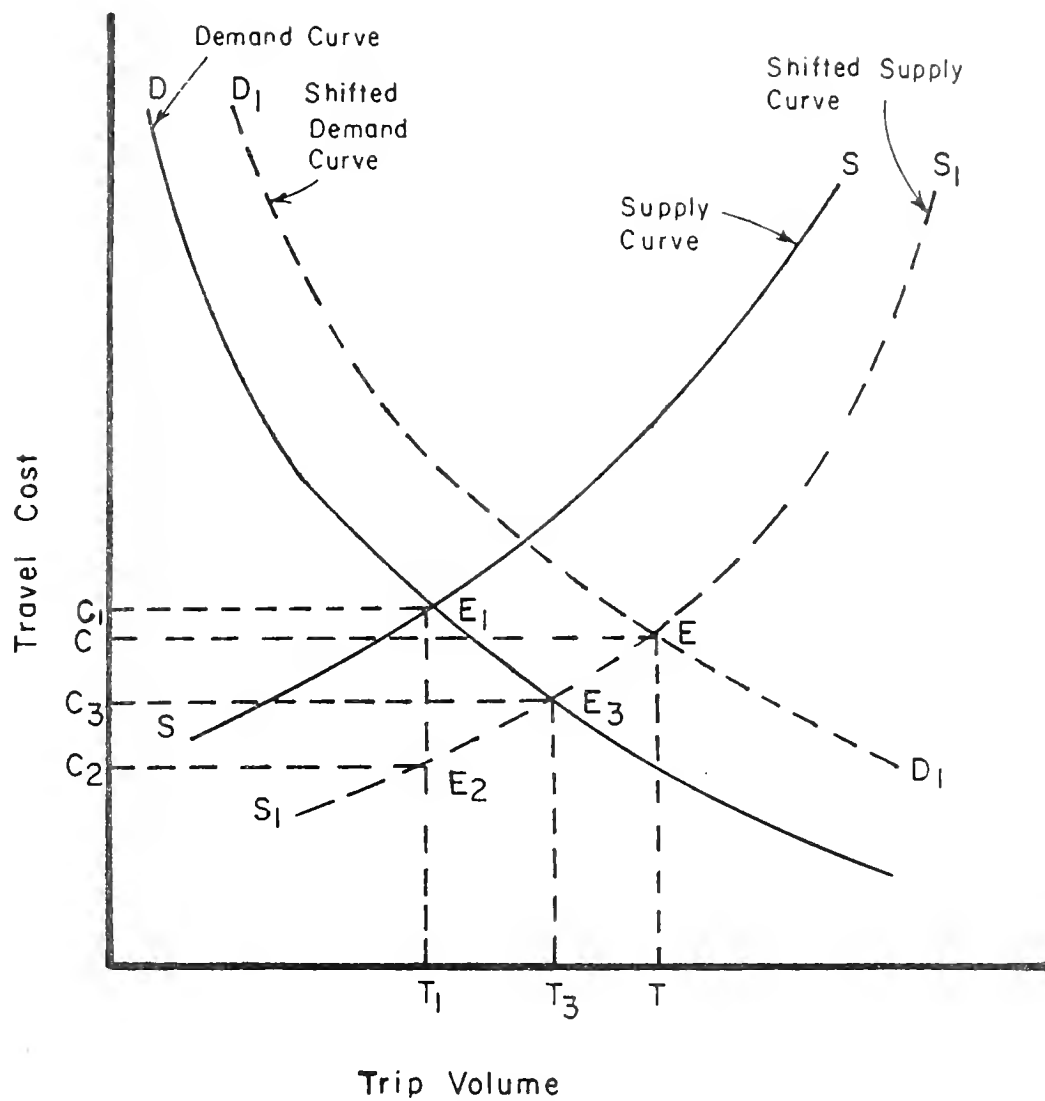


Figure 5. Diagrammatic Presentation of Transportation Supply and Demand Theory



Table 1. Assumptions Used in Development of the Path Model Describing the Relationship of Transportation Development and City Development

No.	Description of Assumption
1	City location and inter-city transportation system influence intra-city transportation development.
2	Intra-city transportation development influence industrial development.
3	Industrial development influences employment.
4	Employment influences mean family income.
5	Mean family income influences total population.
6	City population influences commercial land use.
7	Vacant land use influences institutional land use.
8	Institutional land use influences residential land use.



but come from the author's understanding of the cities growth process; although assumptions three to five are close to assumptions made in "Lowery" type urban development models which assume basic employment in the long run determines city population (17).

The path model based on the assumptions of Table 1 is shown in Figures 6a and 6b. The conventions used in these diagrams are the same as those explained in conjunction with Figures 4a and 4b. Figure 6a which corresponds with Figure 4a shows assumed relationship between the actual variables and the associated path regression coefficients; while Figure 6b shows the normalized variables and the path coefficients. The original form of the path model is more convenient for the verification of the assumed links and prediction of various land uses; while the second form of the model is more convenient for calculation of the overall effects of one variable on another.

The operational form of the path model corresponding with the path diagram of Figure 6a is shown by the nine sequential regression equations 6a to 6h shown in Figure 7a. The coefficient of these equations then are path regression coefficients which upon verification are used to predict the endogenous variables. The mathematical form of the path model with normalized coefficients and path coefficients is presented by sequential regression equations 7a to 7h, shown in Figure 7b. After verification, these path coefficients are used to determine the overall effect of causal variables on dependent variables. If any of the path coefficients prove to be non-significant, then the corresponding link on the path diagrams of Figures 6a and 6b is erased and the resulting new model analyzed for further verification.





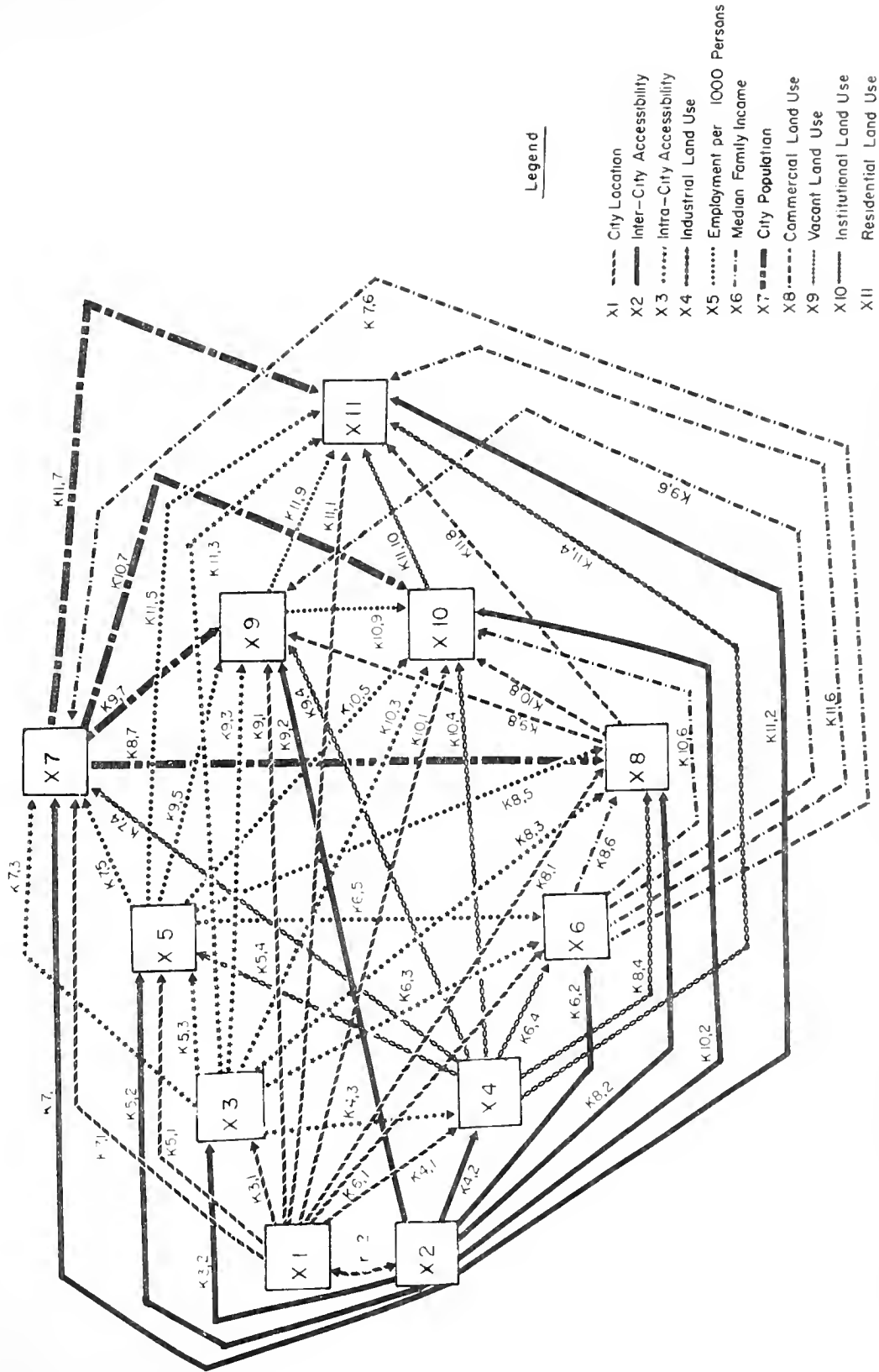


FIGURE 6a TRANSPORTATION CITY DEVELOPMENT PATH MODEL WITH UNNORMALIZED VARIABLES AND PATH REGRESSION COEFFICIENTS



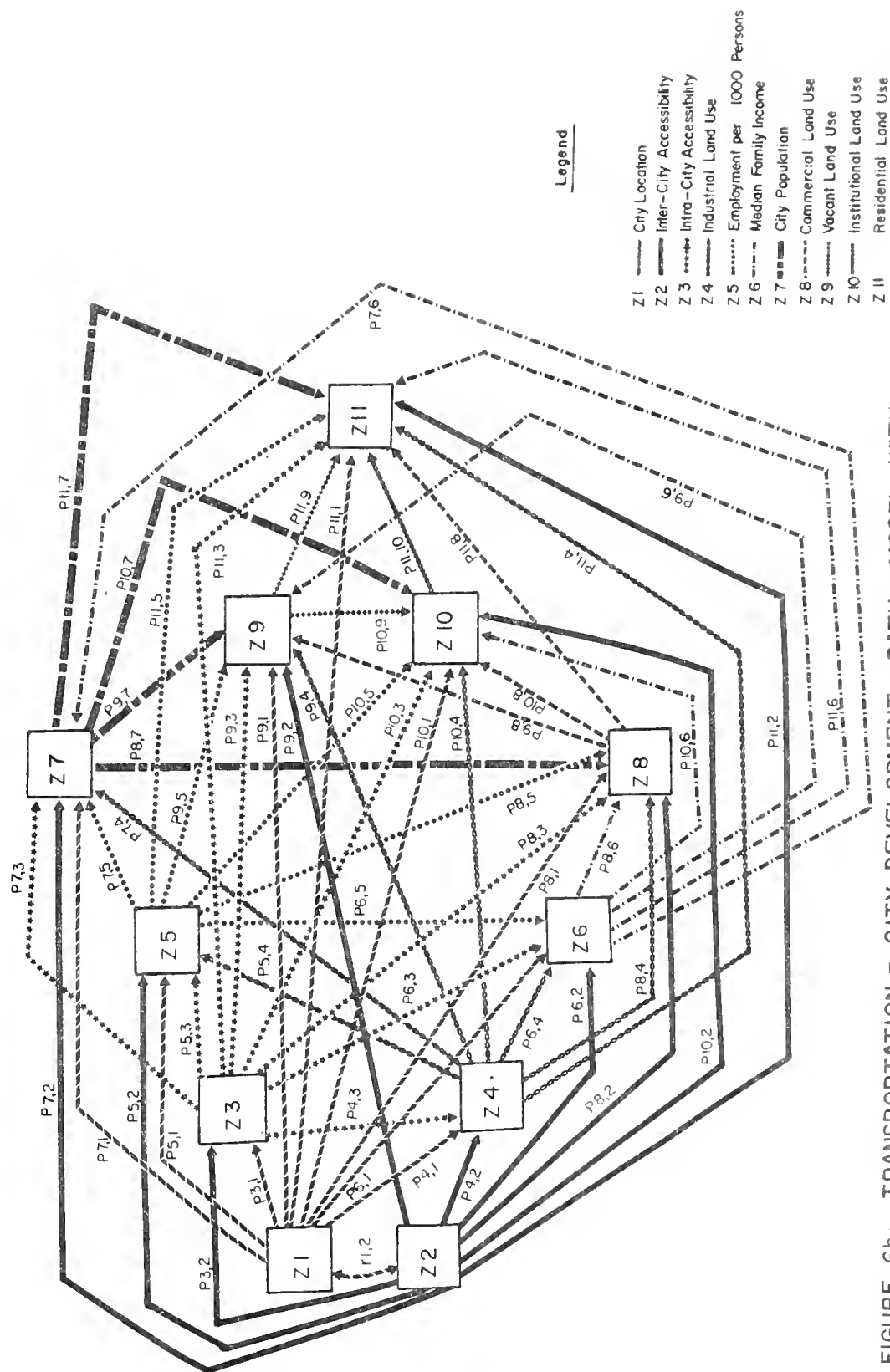


FIGURE 6b. TRANSPORTATION - CITY DEVELOPMENT PATH MODEL WITH NORMALIZED VARIABLES AND PATH COEFFICIENTS



$X_3 = K_{3,1} X_1 + K_{3,2} X_2 + K_{3,e} E_3$	Equation 6a
$X_4 = K_{4,1} X_1 + K_{4,2} X_2 + K_{4,3} X_3 + K_{4,4} X_4 + K_{4,e} E_4$	Equation 6b
$X_5 = K_{5,1} X_1 + K_{5,2} X_2 + K_{5,3} X_3 + K_{5,4} X_4 + K_{5,e} E_5$	Equation 6c
$X_6 = K_{6,1} X_1 + K_{6,2} X_2 + K_{6,3} X_3 + K_{6,4} X_4 + K_{6,5} X_5 + K_{6,e} E_6$	Equation 6d
$X_7 = K_{7,1} X_1 + K_{7,2} X_2 + K_{7,3} X_3 + K_{7,4} X_4 + K_{7,5} X_5 + K_{7,6} X_6 + K_{7,e} E_7$	Equation 6e
$X_8 = K_{8,1} X_1 + K_{8,2} X_2 + K_{8,3} X_3 + K_{8,4} X_4 + K_{8,5} X_5 + K_{8,6} X_6 + K_{8,7} X_7 + K_{8,e} E_8$	Equation 6f
$X_9 = K_{9,1} X_1 + K_{9,2} X_2 + K_{9,3} X_3 + K_{9,4} X_4 + K_{9,5} X_5 + K_{9,6} X_6 + K_{9,7} X_7 + K_{9,8} X_8 + K_{9,e} E_9$	Equation 6g
$X_{10} = K_{10,1} X_1 + K_{10,2} X_2 + K_{10,3} X_3 + K_{10,4} X_4 + K_{10,5} X_5 + K_{10,6} X_6 + K_{10,7} X_7 + K_{10,8} X_8 + K_{10,9} X_9 + K_{10,e} E_{10}$	Equation 6h
$X_{11} = K_{11,1} X_1 + K_{11,2} X_2 + K_{11,3} X_3 + K_{11,4} X_4 + K_{11,5} X_5 + K_{11,6} X_6 + K_{11,7} X_7 + K_{11,8} X_8 + K_{11,9} X_9 + K_{11,10} X_{10} + K_{11,e} E_{11}$	Equation 6i

Figure 7a. Path Model with Unnormalized Variables and Regression Coefficients

$Z_3 = P_{3,1} Z_1 + P_{3,2} Z_2 + P_{3,e} E_3$	Equation 7a
$Z_4 = P_{4,1} Z_1 + P_{4,2} Z_2 + P_{4,3} Z_3 + P_{4,e} E_4$	Equation 7b
$Z_5 = P_{5,1} Z_1 + P_{5,2} Z_2 + P_{5,3} Z_3 + P_{5,4} Z_4 + P_{5,e} E_5$	Equation 7c
$Z_6 = P_{6,1} Z_1 + P_{6,2} Z_2 + P_{6,3} Z_3 + P_{6,4} Z_4 + P_{6,5} Z_5 + P_{6,e} E_6$	Equation 7d
$Z_7 = P_{7,1} Z_1 + P_{7,2} Z_2 + P_{7,3} Z_3 + P_{7,4} Z_4 + P_{7,5} Z_5 + P_{7,6} Z_6 + P_{7,e} E_7$	Equation 7e
$Z_8 = P_{8,1} Z_1 + P_{8,2} Z_2 + P_{8,3} Z_3 + P_{8,4} Z_4 + P_{8,5} Z_5 + P_{8,6} Z_6 + P_{8,7} Z_7 + P_{8,e} E_8$	Equation 7f
$Z_9 = P_{9,1} Z_1 + P_{9,2} Z_2 + P_{9,3} Z_3 + P_{9,4} Z_4 + P_{9,5} Z_5 + P_{9,6} Z_6 + P_{9,7} Z_7 + P_{9,8} Z_8 + P_{9,e} E_9$	Equation 7g
$Z_{10} = P_{10,1} Z_1 + P_{10,2} Z_2 + P_{10,3} Z_3 + P_{10,4} Z_4 + P_{10,5} Z_5 + P_{10,6} Z_6 + P_{10,7} Z_7 + P_{10,8} Z_8 + P_{10,9} Z_9 + P_{10,e} E_{10}$	Equation 7h
$Z_{11} = P_{11,1} Z_1 + P_{11,2} Z_2 + P_{11,3} Z_3 + P_{11,4} Z_4 + P_{11,5} Z_5 + P_{11,6} Z_6 + P_{11,7} Z_7 + P_{11,8} Z_8 + P_{11,9} Z_9 + P_{11,10} Z_{10} + P_{11,e} E_{11}$	Equation 7i

Figure 7b. Path Model with Normalized Variable and Path Coefficients



Table 2. Designation and Names of Variables Used in the Transportation-Urban Development Path Model

Variable Designation	Variable Name
$X_1$	City location *
$X_2$	Inter-city accessibility *
$X_3$	Intra-city accessibility *
$X_4$	Industrial land use **
$X_5$	Employment per 1000 persons *
$X_6$	Median family income *
$X_7$	City population *
$X_8$	Commercial land use **
$X_9$	Vacant land use **
$X_{10}$	Institutional land use **
$X_{11}$	Residential land use **

\* exogenous variables

\*\*endogenous variables





The transportation-urban development path model contains 11 variables. These variables are listed in Table 2 for purposes of identification of model variables. Detailed definitions of these variables are sensitive to the specific data measurement techniques. These 11 variables as used in this study are further defined in the data collection section.

It should be noted at this time that these variables fall into two categories of endogenous and exogenous variables. It is the purpose of the model to explain the variations of endogenous variables; while the values of the exogenous variables are assumed to be known. The purpose of inclusion of the exogenous variables in the model is to assist in explanation of the variations of the endogenous variables. The purpose of this model is to explain variations of various land uses for a given socio-economic and transportation system. Therefore residential land use, institutional land use, vacant land use, commercial land use, and industrial land use are the real dependent (endogenous) variables and the city location inter-city accessibility, intra-city accessibility, employment rate, median family income, and population are really the set of independent (exogenous) variables to assist in the explanation of variations of the dependent variables.

#### References

1. Charles River Associates, Inc., Measurement of the Effects of Transportation Changes, prepared for DOT, Sept. 1972, Chapter 4, p. 40.
2. Campbell, D. T., and Stanley, J. C., Experimental and Quasi-Experimental Designs for Research, Rand McNally and Company, 1966.



3. Wold, Herman, "Causal Inference from Observational Data: A Review of Ends and Means", Journal of Royal Statistical Society, Sec. A., 1956, pp. 28-50.
4. Ibid, pp. 28-32.
5. Ibid, p. 29.
6. Ibid.
7. Ibid., p. 31.
8. Ibid.
9. Ibid.
10. Ibid, p. 35.
11. Land, Kenneth, "Principles of Path Analysis", pp. 3-37 in Edgar Borgetta (ed.), Sociological Methodology, Jossey Bass, San Francisco, 1969.
12. Ibid., p. 5.
13. Nie, Norman H., et al, Statistical Package for the Social Sciences, SPSS, Second Edition, McGraw-Hill, New York, 1975, pp. 383-396.
14. Land, Kenneth, pp. 6-7.
15. Ibid., p. 7.
16. Hutchinson, B. G., Principle of Urban Transport System Planning, McGraw-Hill, New York, 1974, pp. 306-325.
17. Moore, T. W., et al, An Introduction to Urban Development Models and Guidelines for their Use in Transportation Planning, U.S. Department of Transportation, Oct. 1975, p. 69.



## CHAPTER FOUR

### DESIGN OF QUASI-EXPERIMENT

The design of this quasi-experiment involves observations and statistical analysis of four series of cross-sectional data to test a hypothesized relationship between transportation development and urban development. In a traditional design of experiment the significance of a treatment is obtained by subjecting only some of the experimental units to a given treatment and then comparing the experimental units subjected to the treatment with the experimental units not subjected to the treatment to test for significance of the particular treatment. In the quasi-experimental design such control of treatment and experimental units is not possible, but the control is obtained by randomly sampling an experimental unit and then measuring the unique treatments observed within the sampled experimental unit. This sampling method is very useful for perspective studies of urban phenomenon because it does not require a prior categorization of cities as to different treatment, but allows for observations of actual treatments after random selection of the city as the experimental unit. The method still allows stratification of sample universe into data cells for other than non-treatment variables. The sampling scheme for choice of cities to be observed for one cell thru time is shown in Figure 8. The nomenclature used in this figure is that of Stanley and Campbell (1).



1940	{	R	$X_A$	$O_A$
		R	$X_B$	$O_B$
		.	.	.
		.	.	.
		.	.	.
		R	$X_N$	$O_N$
1950	{	R	$X_A$	$O_A$
		R	$X_B$	$O_B$
		.	.	.
		.	.	.
		.	.	.
		R	$X_N$	$O_N$
1960	{	R	$X_A$	$O_A$
		R	$X_B$	$O_B$
		.	.	.
		.	.	.
		.	.	.
		R	$X_N$	$O_N$
1970	{	R	$X_A$	$O_A$
		R	$X_B$	$O_B$
		.	.	.
		.	.	.
		.	.	.
		R	$X_N$	$O_N$

Figure 8. Sampling Scheme for Calibration of Transportation-City Development Path Model





Each line of this figure represents the process by which one city is selected and observed. Letter R signifies the randomness of the process. Letter  $X_A$  signifies that this city is subjected to a unique set of treatments ( $X_A$ ), and  $O_A$  signifies that the particular treatments,  $X_A$ , are observed and measured.

#### Procedure to Select Sample Cities

The "quasi-experimental" design chosen to obtain observed data to calibrate the transportation-urban development path model of equations 6a to 6h is "cross-sectional" sampling. Samples of all Indiana cities in the inference space, in random within strata of thirty-six, are selected, and then values of the eleven variables in the path model for each of the sample cities are measured.

As a means of obtaining cross-sectional samples of cities on the inference space, the cities are stratified across four time periods, three geographic locations, and three population sizes. The resulting 36 cells; nine in each of the four time blocks are shown in Figure 9. Random sampling of cities in each of these 36 cells and aggregation of the selected cities in each of the nine cells of a time block results in the four lists of independent cross-sectional samples.

There were several reasons for sampling cities in 1940, 1950, 1960 and 1970 periods. The first sampling period of 1940 was chosen because for smaller cities some socio-economic variables first became available in the 1940 census publications. Also, 1940 was prior to the accelerated move of city dwellers to suburbia, as well as relatively heavy demand on and investment for the transportation system brought about by the second



Time	City Location	City Population		
		2,500 - 10,000	10,001 - 20,000	20,001 - 50,000
1940	North	Cell 1	Cell 4	Cell 7
	Central	Cell 2	Cell 5	Cell 8
	South	Cell 3	Cell 6	Cell 9
1950	North	Cell 10	Cell 11	Cell 12
	Central	Cell 13	Cell 14	Cell 15
	South	Cell 16	Cell 17	Cell 18
1960	North	Cell 19	Cell 20	Cell 21
	Central	Cell 22	Cell 23	Cell 24
	South	Cell 25	Cell 26	Cell 27
1970	North	Cell 28	Cell 29	Cell 30
	Central	Cell 31	Cell 32	Cell 33
	South	Cell 34	Cell 35	Cell 36

Figure 9. Stratification of Indiana Cities Across Time, Geographic Location, and Population Size



world war. Choice of decadal years was to enable direct use of Bureau of Census decadal socio-economic data in the model calibrations.

The choice of 1970 as the last sampling period was due to difficulty of obtaining uniform data for any later date; as well as the desire for the same time differences between the sampling periods.

Because of the assumption that city location was a factor in the city development the State of Indiana was stratified into three zones to assure representative city samples from different geographic locations. Northern Indiana was prime agricultural land and manufacturing more so than the more hilly southern section. Generally it is assumed that the dividing line between north and south is U.S. 40 running thru Indianapolis. But it was decided to choose an intermediate central zone to provide a gradual transition between the northern and southern zone. Also, central Indiana is unique because it includes Indianapolis which is the state's largest city and a nationally unique transportation hub for a city of its size; in part because of the four interstate highways passing thru the city. The exact boundary of the three locational zones follow the boundary lines of planning and development regions published by the Indiana State Planning Services Agency (2) as shown in Figure 10. Aggregation of several counties into a development region implies that the counties have in common similar resources and planning goals. If the city location in this experiment proved significant, then each of the three geographical zones could be further disaggregated to the level of Indiana planning and development regions.



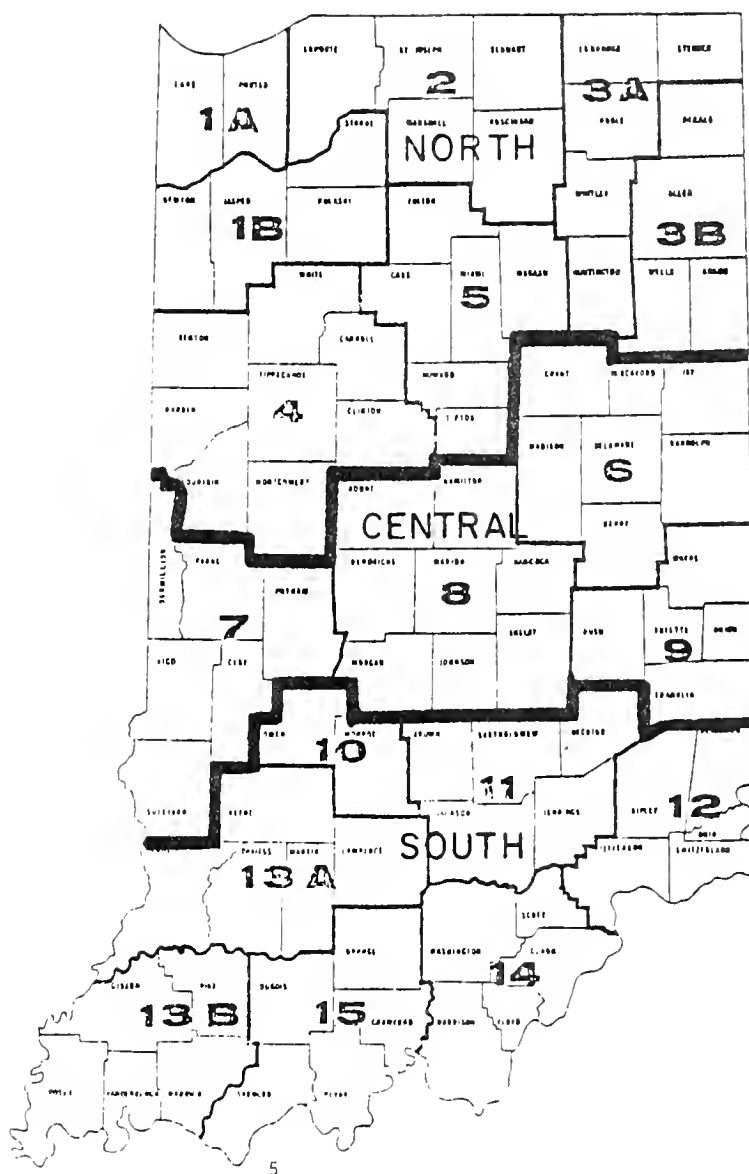


Figure 10. Division of State of Indiana Into Three Zones

Source for the base map showing Indiana Planning and Development Regions: (Indiana) State Planning Services Agency, "A Directory of Indiana Local and Regional Planning and Development Commissions", Indianapolis, June 1975, p. 5.





The sample space was further stratified in accordance to the cities population size. It has been already explained that the scope of the experiment was limited to cities with less than 50,000 population (not located in any SMSA county) and greater than 2,500. The cities between 2,500 and 50,000 were further stratified to three levels of 2,500-10,000; 10,001-20,000; 20,001-50,000. Stratification according to population was necessary because there were many more smaller cities than the larger ones and without stratification an unbiased cross-sectional sample could not be obtained. The particular boundary line of 20,000 was chosen to assure a minimum of one city in each cell.

The specific sampling procedure required identification of all cities in the inference space and listing them in their proper cell prior to random sampling. As a first step in determining the cities in the inference space all cities within Standard Metropolitan Statistical Areas (SMSA) were omitted from the sample space to conform to the scope of the project. The excluded and eligible counties for the four sampling periods are shown in Figures 11 to 14 adopted from census publications (3,4,5,6).

It is noted that the Census Bureau in 1940 did not designate counties as metropolitan areas, but the cities of South Bend, Fort Wayne, Indianapolis, Terre Haute, and Evansville were designated as metropolitan areas (7) and for consistency all cities in the counties containing these metropolitan areas were excluded from this study.

The eligible, non-SMSA, counties in each of the Figures 11 to 14 were selected one at a time and, by use of the Bureau of Census data, (3,4,5,6) all the cities with greater than 2,500 populations were



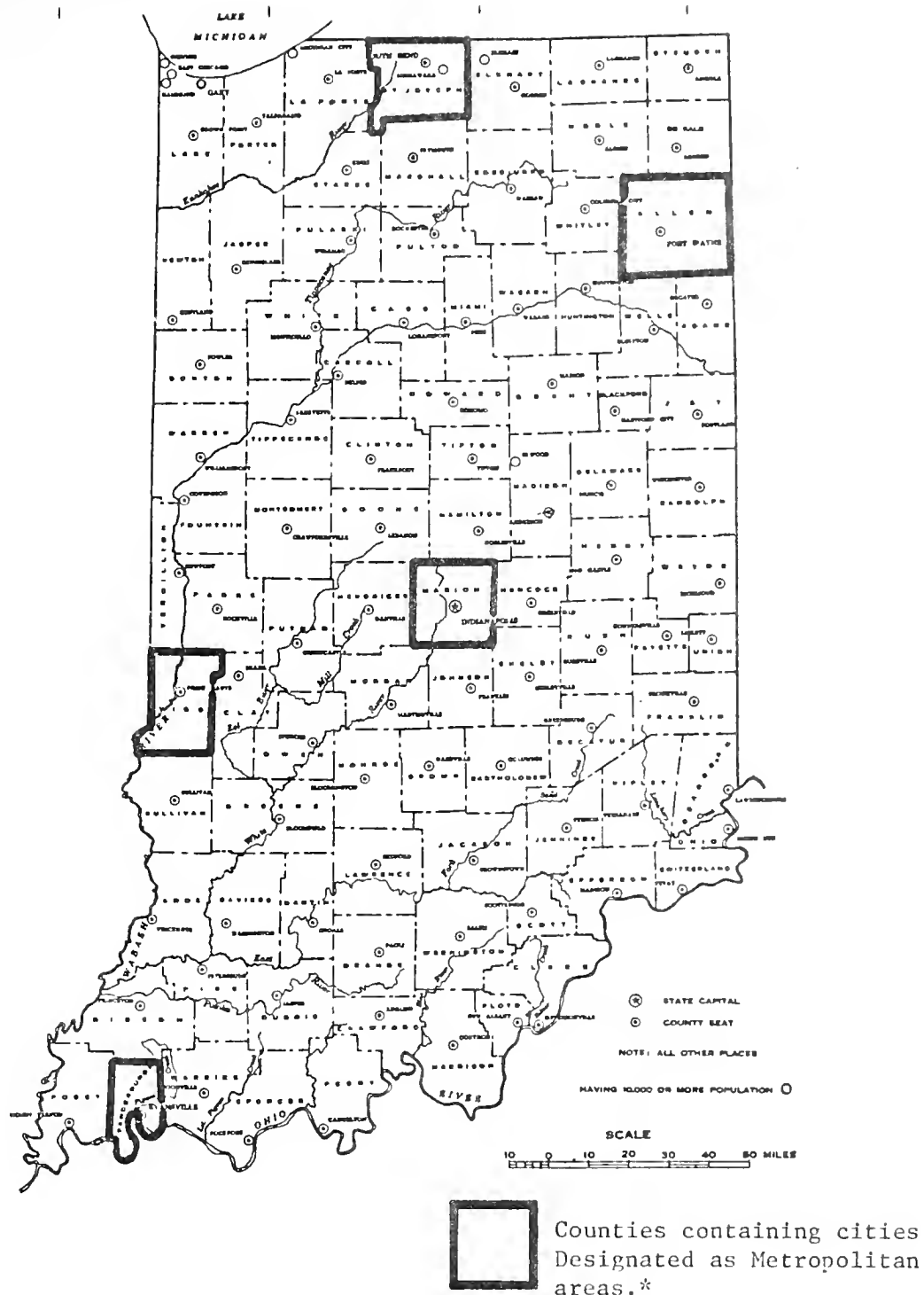


Figure 11. Counties Excluded (Containing Cities Designated as Metropolitan Areas) from 1940 Data

Source: U.S. Bureau of Census. Census of Population: 1940, Vol. II, Characteristics of the Population, Part 2 Florida-Iowa. U.S. Government Printing Office, Washington, D.C., p. 674 (Base map) and pp. 842-844 (Designation of Metropolitan Areas)\*.



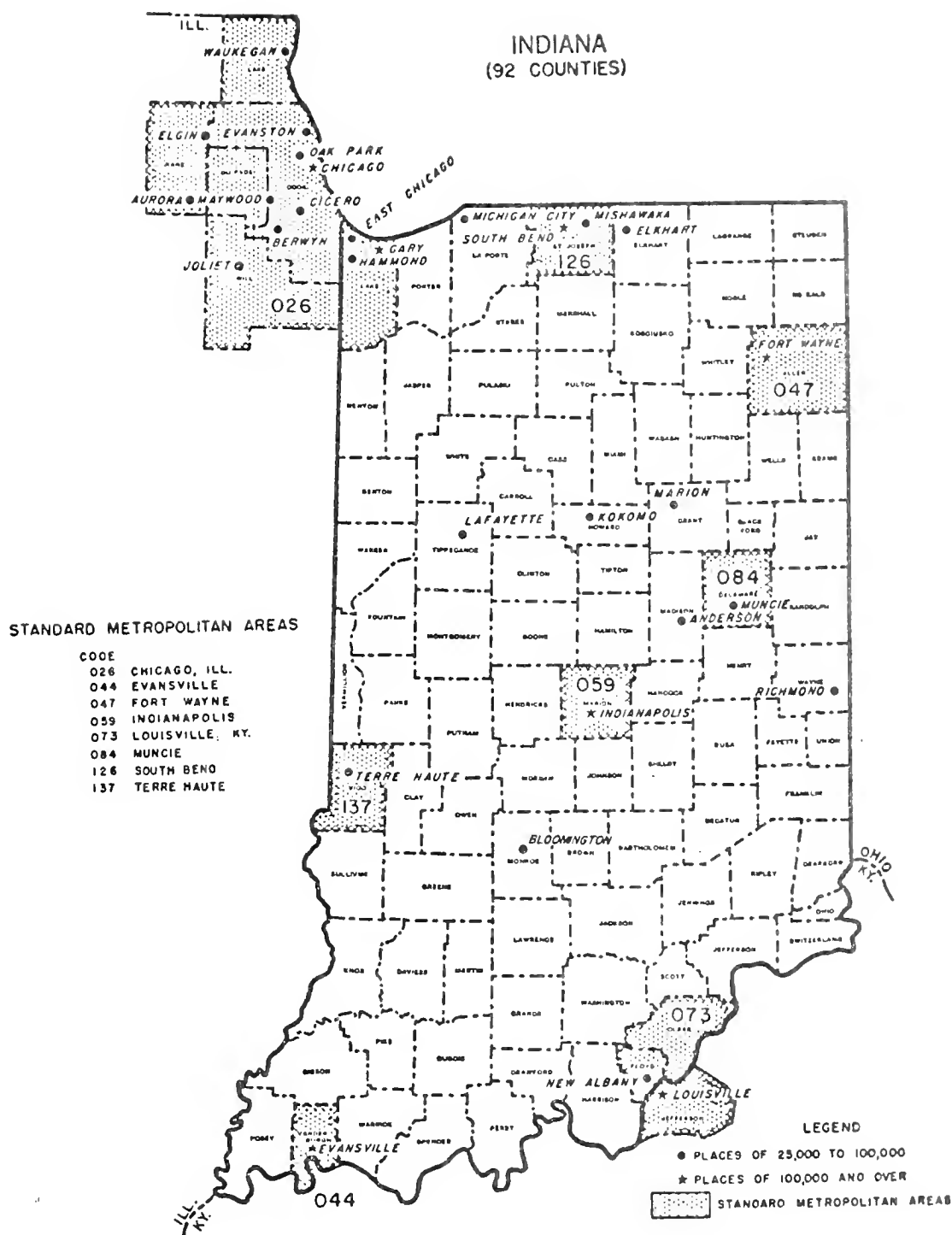


Figure 12. Counties Excluded (Standard Metropolitan Areas) from 1950 Data

Source: U.S. Bureau of Census. U.S. Census of Housing: 1950 Vol. I, General Characteristics, Part 3, Idaho-Massachusetts, U.S. Government Printing Office, Washington, D. C., 1953.



# INDIANA-- COUNTIES, PLACES OF 25,000 OR MORE, AND STANDARD METROPOLITAN STATISTICAL AREAS

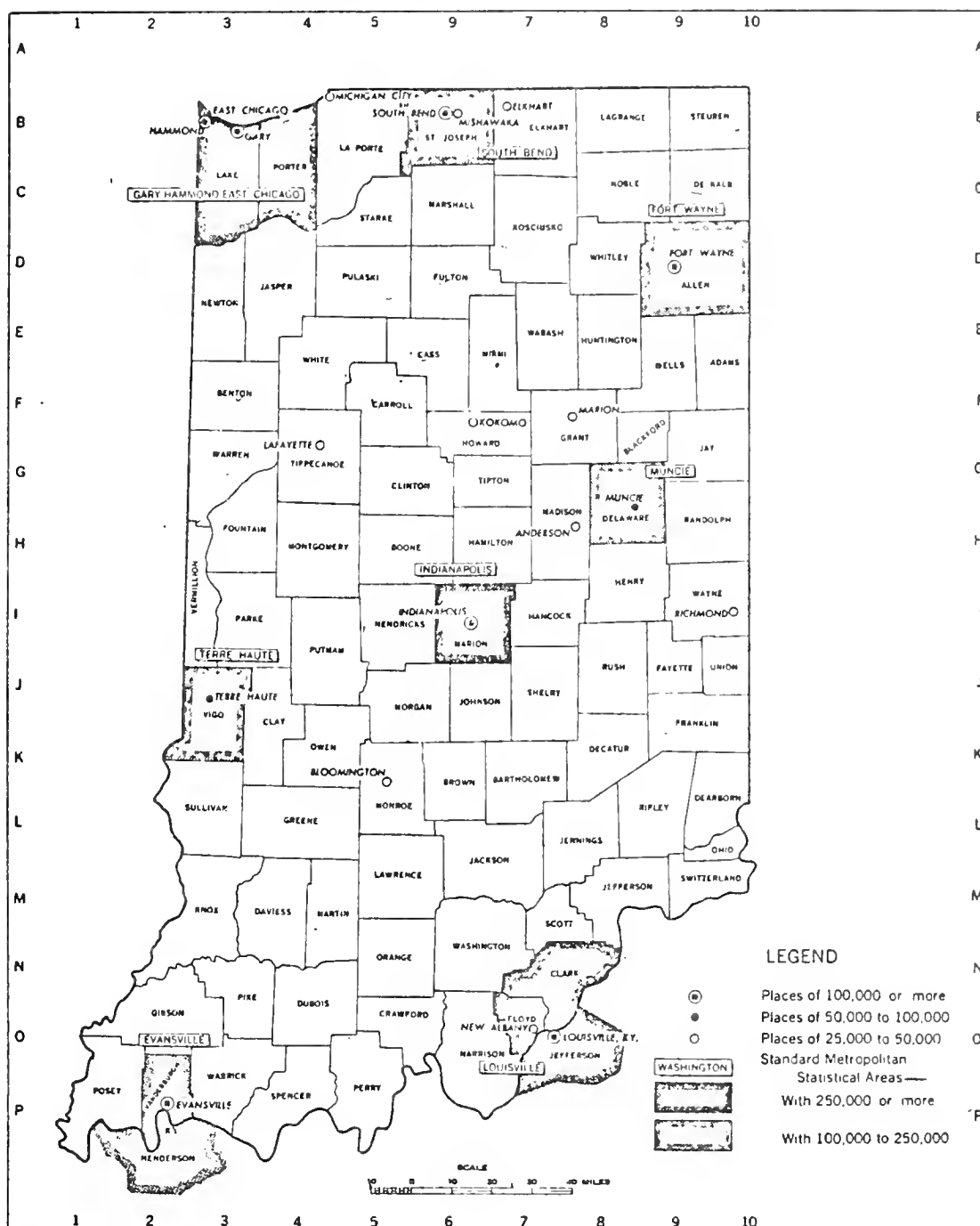


Figure 13. Counties Excluded (Standard Metropolitan Statistical Areas) from 1960 Data

Source: U.S. Bureau of Census. U.S. Census of Population: 1960, Vol. I, Characteristics of the Population. Part 16, Indiana. U.S. Government Printing Office, Washington, D. C., 1963.





## INDIANA

## Counties, Standard Metropolitan Statistical Areas, and Selected Places



Figure 14. Counties Excluded (Standard Metropolitan Statistical Areas) from 1970 Data

Source: U.S. Bureau of Census. U.S. Census of Population: 1970, Vol. I, Characteristics of Population. Part 16, Indiana, U.S. Government Printing Office, Washington, D. C., April 1973.



listed in one of the three population levels as shown in Figure 9. The lists of cities in each county were aggregated to the level of one of the 18 Indiana planning and development regions shown in Figure 10. This in turn allowed further aggregation of the cities in the sample space to the level of one of the three geographical locations for each of the four time periods. Each of the Tables 3 to 14 represent the results of one such aggregation containing three cells. After a decision as to the percentage of sample space to be studied each of the 36 cells in these tables were randomly sampled to yield the list of cities to be investigated.

The appropriate percent of population to be sampled depends on variance of variables to be measured. The variance of the 11 variables of interest in this study were not known. In such a situation, multi-stage experimentation normally was thought to be helpful, but in this case the potential for instrumentation error was felt to be sufficient to encourage a single stage sampling. Table 15 which is based on Tables 3 to 14 shows the number of cities in the sample space for each of the 36 cells. There are a total of 363 cities in the sample space. Based on the project's resources and the desire for a satisfactory single stage sampling; an overall 25 percent sample was chosen. But each of the cells were not sampled equally, because there were many more smaller cities than larger cities, and with equal percent sampling of each cell the experiment would have been biased against larger cities.



Table 3. List of 1940 Cities in the Sample Space - Northern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
1A	Valparaiso			1
1B	Rensselaer			2
2	Plymouth			3
2		LaPorte		4
2			Michigan City	5
2	Warsaw			6
2			Elkhart	7
2		Goshen		8
2	Napanee			9
3A	Columbia City			10
3A	Angola			11
3A	Kendallville			12
3A		Huntington		13
3B	Decatur			14
3B	Auburn			15
3B	Garrett			16
3B	Bluffton			17
4	Monticello			18
4			Lafayette	19
4	W. Lafayette			20
4		Crawfordsville		21
4	Attica			22
4		Frankfort		23
5	N. Manchester			24
5	Wabash			25
5	Tipton			26
5		Peru		27
5			Kokomo	28
5	Rochester			29
5			Logansport	30



Table 4. List of 1940 Cities in the Sample Space - Central Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
6	Union City			1
6	Winchester			2
6	Alexandria			3
6			Anderson	4
6		Elwood		5
6	Dunkirk			6
6	Portland			7
6		New Castle		8
6	Fairmount			9
6			Marion	10
6			Muncie	11
6	Hartford City			12
7	Clinton			13
7	Sullivan			14
7	Green Castle			15
7	Brazil			16
8		Shelbyville		17
8	Martinsville			18
8	Franklin			19
8	Noblesville			20
8	Greenfield			21
8	Lebanon			22
9			Richmond	23
9	Rushville			24
9		Connersville		25





Table 5. List of 1940 Cities in the Sample Space - Southern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
10			Bloomington	1
11	North Vernon			2
11	Seymour			3
11	Greensburg			4
11		Columbus		5
12	Batesville			6
12	Madison			7
12	Aurora			8
12	Lawrenceburg			9
13A		Bedford		10
13A	Mitchell			11
13A	Bicknell			12
13A		Vincennes		13
13A	Jasonville			14
13A	Linton			15
13A	Washington			16
13B	Booneville			17
13B	Oakland			18
13B	Princetown			19
13B	Petersburg			20
14	Mt. Vernon			21
15	Salem			22
15	Tell City			23
15	Huntingburg			24
15	Jasper			25



Table 6. List of 1950 Cities in the Sample Space - Northern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
1A	Chesterton			1
1A		Valparaiso		2
1B	Knox			3
1B	Rensselaer			4
2	Bremen			5
2	Plymouth			6
2		LaPorte		7
2			Michigan City	8
2	Warsaw			9
2			Elkhart	10
2		Goshen		11
2	Napaneer			12
3A	Columbia City			13
3A	Angola			14
3A	Kendallville			15
3A		Huntington		16
3B	Decatur			17
3B	Auburn			18
3B	Garrett			19
3B	Bluffton			20
4	Monticello			21
4			Lafayette	22
4		W. Lafayette		23
4		Crawfordsville		24
4	Delphi			25
4	Attica			26
4		Frankfort		27
5	N. Manchester			28
5		Wabash		29
5	Tipton			30
5		Peru		31
5			Kokomo	32
5	Rochester			33
5			Logansport	34



Table 7. List of 1950 Cities in the Sample Space - Central Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
6	Union City			1
6	Winchester			2
6	Alexandria			3
6			Anderson	4
6		Elwood		5
6	Dunkirk			6
6	Portland			7
6		New Castle		8
6	Fairmount			9
6	Gas City			10
6			Marion	11
6	Hartford City			12
7	Clinton			13
7	Sullivan			14
7	Greencastle			15
7	Brazil			16
8		Shelbyville		17
8	Martinsville			18
8	Edinburg			19
8	Franklin			20
8	Greenwood			21
8	Noblesville			22
8	Greenfield			23
8	Danville			24
8	Plainfield			25
8	Lebanon			26
9	Cambridge City			27
9	Rushville			28
9			Richmond	29
9	Brookville			30
9		Connersville		31



Table 8. List of 1950 Cities in the Sample Space - Southern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
10			Bloomington	1
11	North Vernon			2
11	Seymour			3
11	Greensburg			4
11		Columbus		5
12	Batesville			6
12	Madison			7
12	Aurora			8
12	Lawrenceburg			9
13A		Bedford		10
13A	Mitchell			11
13A	Bicknell			12
13A		Vincennes		13
13A	Jasonville			14
13A	Linton			15
13A		Washington		16
13B	Booneville			17
13B	Oakland City			18
13B	Princetown			19
13B	Petersburg			20
13B	Mt. Vernon			21
14	Salem			22
14	Austin			23
14	Scottsburg			24
15	Rockport			25
15	Tell City			26
15	Paoli			27
15	Huntingburg			28
15	Jasper			29





Table 9. List of 1960 Cities in the Sample Space - Northern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
1B	Knox			1
1B	Rensselaer			2
2	Bremen			3
2	Plymouth			4
2			LaPorte	5
2			Michigan City	6
2	Warsaw			7
2			Elkhart	8
2		Goshen		9
2	Napanee			10
3A	Columbia City			11
3A	Angola			12
3A	Kendallville			13
3A	Ligonier			14
3A		Huntington		15
3B	Berne			16
3B	Decatur			17
3B	Auburn			18
3B	Carrett			19
3B	Bluffton			20
4	Monticello			21
4			Lafayette	22
4		W. Lafayette		23
4		Crawfordsville		24
4	Delphi			25
4	Attica			26
4	Covington			27
4		Frankfort		28
5	N. Manchester			29
5		Wabash		30
5	Tipton			31
5		Peru		32
5			Kokomo	33
5	Rochester			34
5			Logansport	35



Table 10. List of 1960 Cities in the Sample Space - Central Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
6	Union City			1
6	Winchester			2
6	Alexandria			3
6			Anderson	4
6	Chesterfield			5
6		Elwood		6
6	Dunkirk			7
6	Portland			8
6			New Castle	9
6	Fairmount			10
6	Gas City			11
6			Marion	12
6	Hartford City			13
7	Clinton			14
7	Sullivan			15
7	Greencastle			16
7	Rockville			17
7	Brazil			18
8		Shelbyville		19
8	Martinsville			20
8	Mooresville			21
8	Edinburg			22
8	Franklin			23
8	Greenwood			24
8	New Whiteland			25
8	Noblesville			26
8	Greenfield			27
8	Brownsburg			28
8	Danville			29
8	Plainfield			30
8	Lebanon			31
9	Cambridge City			32
9			Richmond	33
9	Rushville			34
9	Brookville			35
9		Connersville		36



Table 11. List of 1960 Cities in the Sample Space - Southern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
10			Bloomington	1
10	Spencer			2
11	North Vernon			3
11		Seymour		4
11	Greensburg			5
11			Columbus	6
12	Batesville			7
12		Madison		8
12	Aurora			9
12	Lawrenceburg			10
13A	Loogootee			11
13A		Bedford		12
13A	Mitchell			13
13A	Bicknell			14
13A		Vincennes		15
13A	Linton			16
13A		Washington		17
13B	Booneville			18
13B	Oakland City			19
13B	Princetown			20
13B	Petersburg			21
13B	Mt. Vernon			22
14	Salem			23
14	Austin			24
14	Scottsburg			25
14	Corydon			26
15	Rockport			27
15	Tell City			28
15	Paoli			29
15	Huntingburg			30
15	Jasper			31



Table 12. List of 1970 Cities in the Sample Space - Northern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
1B	Knox			1
1B	Rensselaer			2
2	Bremen			3
2	Plymouth			4
2			LaPorte	5
2	Long Beach			6
2			Michigan City	7
2	Trial Creek			8
2	Westville			9
2	Warsaw			10
2	Winona Lake			11
2			Elkhart	12
2		Goshen		13
3A	Napanea			14
3A	Columbia City			15
3A	Angola			16
3A	Kendallville			17
3A	Ligonier			18
3B		Huntington		19
3B	Berne			20
3B	Decatur			21
3B	Auburn			22
3B	Garrett			23
4	Bluffton			24
4	Monticello			25
4		Crawfordsville		26
4	Delphi			27
4	Fowler			28
4	Attica			29
4	Covington			30
5		Frankfort		31
5	N. Manchester			32
5		Wabash		33
5	Tipton			34
5		Peru		35
5			Kokomo	36
5	Rochester			37





Table 13. List of 1970 Cities in the Sample Space - Central Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
6	Union City			1
6	Winchester			2
6	Alexandria			3
6	Dunkirk			4
6	Portland			5
6			New Castle	6
6	Fairmount			7
6	Gas City			8
6			Marion	9
6	Upland			10
6	Hartford City			11
7	Greencastle			12
7	Rockville			13
9	Cambridge City			14
9			Richmond	15
9	Rushville			16
9	Brookville			17
9		Connersville		18



Table 14. List of 1970 Cities in the Sample Space - Southern Zone

Urban Planning and Devel- opment Region*	City Population			No.
	2,500-10,000	10,001-20,000	20,001-50,000	
10			Bloomington	1
11	Mt. Vernon			2
11		Seymour		3
11	Greensburg			4
11			Columbus	5
12	Batesville			6
12	Hanover			7
12		Madison		8
12	Aurora			9
12	Lawrenceburg			10
13A	Loogootee			11
13A		Bedford		12
13A	Mitchell			13
13A	Bicknell			14
13A		Vincennes		15
13A	Bloomfield			16
13A	Linton			17
13A		Washington		18
13B	Booneville			19
13B	Fort Branch			20
13B	Oakland City			21
13B	Princetown			22
13B	Petersburg			23
13B	Mt. Vernon			24
14	Salem			25
14	Scottsburg			26
14	Corydon			27
15	Rockport			28
15	Tell City			29
15	Paoli			30
15	Huntingburg			31
15	Jasper			32



The cells were sampled according to a maximum criteria of 20 percent and a minimum criteria of two cities. The actual number of cities to be sampled from each cell was included in each of the 36 cells of Table 15. The only exception to the above minimum criteria occurred in three out of the thirty-six cells which contained only one city.

Once the number of cities to be sampled in each of the 36 cells were known, a random number table was used to obtain names of the actual cities to be sampled from each cell among the cities listed in Tables 3 to 12. The sample cities so chosen are listed in Tables 16 to 19. Above tables are ordered to represent all the sample cities for the time periods of 1940, 1950, 1960 or 1970 for the areas shown in Figures 11 to 14. The above scheme resulted in 98 cities to be studied, which means 27 percent of cities in the sample space were sampled.

#### Data Collection

It is necessary to measure the values of the 11 model variables listed in Table 2 for the 98 sample cities listed in Tables 16 to 19 to calibrate the path model of equations 6a to 6i. But prior to actual data collection, the 11 model variables are to be further defined. The names of the variables as given in Table 2 generally convey the basic meaning of each of the variables, but for data collection purposes specific operational definitions of these variables are needed. The operational definitions and measurement units differ from one study to another based on the dimensions of



Table 15. Number of Cities in the Sample Space and Number of Cities Sampled in Each of the Cells

Time	City Location	City Population					
		2,500 - 10,000		10,001 - 20,000		20,001 - 50,000	
		Total	Sample	Total	Sample	Total	Sample
1940	North	19	4	6	2	5	2
	Central	17	3	4	2	4	2
	South	21	4	3	2	1	1
1950	North	20	4	9	2	5	2
	Central	24	5	4	2	3	2
	South	24	5	4	2	1	1
1960	North	22	4	7	2	6	2
	Central	29	6	3	2	4	2
	South	23	5	5	2	2	2
1970	North	26	5	7	2	4	2
	Central	14	3	1	1	3	2
	South	26	5	4	2	3	2
Total Cities in Sample Space 265		57		41			
Total Cities Sampled		53		23		22	

= 363

= 98





Table 16. Cross-Sectional Samples of Non-Metropolitan Indiana Cities for 1940

Location	City Population		
	2,500-10,000	10,001-20,000	20,001-50,000
North	Valparaiso	Huntington	Lafayette
North	Warsaw	Frankfort	Kokomo
North	Kendallville		
North	Attica		
Central	Dunkirk	Shelbyville	Muncie
Central	Hartford City	Connersville	Richmond
Central	Greencastle		
South	Greensburg	Bedford	Bloomington
South	Batesville	Vincennes	
South	Linton		
South	Jasper		

Table 17. Cross-Sectional Samples of Non-Metropolitan Indiana Cities for 1950

Location	City Population		
	2,500-10,000	10,001-20,000	20,001-50,000
North	Napanea	LaPorte	Kokomo
North	Monticello	Frankfort	Logansport
North	Tipton		
North	Rochester		
Central	Hartford City		
Central	Alexandria	New Castle	Anderson
Central	Portland	Shelbyville	Richmond
Central	Greencastle		
Central	Franklin		
South	Aurora	Bedford	Bloomington
South	Bicknell	Vincennes	
South	Petersburg		
South	Tell City		
South	Jasper		



Table 18. Cross-Sectional Samples of Non-Metropolitan Indiana Cities for 1960

Location	City Population		
	2,500-10,000	10,001-20,000	20,001-50,000
North	Plymouth	Goshen	Michigan City
North	Berne	W. Lafayette	Logansport
North	Covington		
North	Tipton		
Central	Chesterfield	Elwood	Anderson
Central	Rockville	Shelbyville	Richmond
Central	Martinsville		
Central	Edinburg		
Central	Noblesville		
Central	Lebanon		
South	Lawrenceburg	Madison	Bloomington
South	Bicknell	Washington	Columbus
South	Booneville		
South	Petersburg		
	Scottsburg		

Table 19. Cross-Sectional Sample of Non-Metropolitan Indiana Cities for 1970

Location	City Population		
	2,500-10,000	10,001-20,000	20,001-50,000
North	Knox	Goshen	Michigan City
North	Long Beach	Crawfordsville	Elkhart
North	Delphi		
North	Fowler		
North	Covington		
Central	Portland	Connersville	New Castle
Central	Upland		Richmond
Central	Hartford City		
South	Bicknell	Madison	Bloomington
South	Bloomfield	Washington	Columbus
South	Booneville		
South	Fort Branch		
South	Jasper		



inference space and level of aggregation. In the following sections the 11 variables are operationally defined and based on the given definition their respective value is measured.

#### City Location (X1)

Indiana cities were stratified to northern, central and southern zones prior to randomized sampling to test for significance of the geographic location. If the three zones proved significant then further disaggregation of each of the three zones could test for significance of the city location variable on the level of each of the 18 Indiana development and planning regions shown in Figure 10. One way to represent the location variable in the model was by the use of the dummy variables (North=00, central=01, and south=10). Dummy variables in the path model are normally presented as continuous variables (8); and this was accomplished by measuring vertical displacement of sample cities in miles relative to the location of Indianapolis at the center of the state. These distances which were scaled as positive north of Indianapolis and negative south of Indianapolis are shown in Table 20. Measured values of locations thus obtained were continuous dummy variables with a horizontal line thru the center of Indianapolis as the reference line; similar results would have been obtained if the northern or southern state boundary would have been used as the reference line.

#### Inter-City Accessibility (X2)

It is believed that city development is significantly affected by its locational advantage in relation to the regional markets. A



Table 20. Vertical Displacement of Sample Cities in Reference to a Horizontal Line Thru the Center of Indianapolis

1940 Sample			1950 Sample			1960 Sample			1970 Sample		
City Name	Dis- tance (mile)		City Name	Dis- tance (mile)		City Name	Dis- tance (mile)		City Name	Dis- tance (mile)	
VALPARAISO	118		ROCHESTER	88		MICHIGAN CITY	134		MICHIGAN CITY	134	
KENDALLVILLE	117		MONTICELLO	67		GOSHEN	125		LONG BEACH	136	
WARSAN	101		LOGANSPORT	67		PLYMOUTH	112		ELKHART CITY	130	
HUNTINGTON	77		KOKOMO	47		LOGANSPORT	67		GOSHEN	125	
KOKOMO	47		HARTFORD	48		BERNE CITY	61		KNOX	104	
HARTFORD CITY	48		PORTLAND	45		LAFAYETTE	44		FOWLER	59	
LAFAYETTE	43		FRANKFORT	34		TIPTON	35		HARTFORD	48	
ATTICA	36		TIPTON	35		ELWOOD	34		PORTLAND	45	
FRANKFORT	34		ALEXANDRIA	33		COVINGTON	26		COVINGTON	26	
MUNCIE	23		ANDERSON	22		ANDERSON	22		CRAWFORDSVILLE	20	
RICHMOND	4		NEW CASTLE	11		CHESTERFIELD	18		UPLAND	48	
GREENCASTLE	-9		RICHMOND	4		LEBANON	18		NEW CASTLE	11	
CONNERSVILLE	-10		GREENCASTLE	-8		NOBLESVILLE	19		RICHMOND	4	
SHELBYVILLE	-18		SHELBYVILLE	-18		RICHMOND	4		CONNERSVILLE	-9	
GREENSBURG	-32		FRANKLIN	-20		SHELBYVILLE	-18		COLUMBUS	-39	
BATESVILLE	-33		BLOOMINGTON	-41		MARTINSVILLE	-26		BLOOMINGTON	-41	
BLOOMINGTON	-41		AURORA	-49		EDINBURG	-30		MADISON	-71	
LINTON	-50		BEDFORD	-62		COLUMBUS	-39		BICKNELL	-68	
BEDFORD	-62		BICKNELL	-68		BLOOMINGTON	-41		WASHINGTON	-77	
VINCENNES	-74		VINCENNES	-74		MADISON	-71		JASPER	-95	
JASPER	-95		PETERSBURG	-87		BICKNELL	-68		FORTBRANCH	-104	
			JASPER	-95		WASHINGTON	-77		BOONVILLE	-118	
			TELL CITY	-125		PETERSBURG	-87		BLOOMFIELD	-50	
			NAPANEE	114		BOONVILLE	-118		DELPHI	56	
						LAWRENCEBURG	-43				
						SCOTTSBURG	-70				





modified form of "Interaction Potential" is used to obtain the indices of location advantage of sample cities. The "Interaction Potential" is defined by Shafran and Wigman (9) as:

$$I.P. (i,m) = \sum_{j=1}^m \frac{P_i \cdot P_j}{t_{ij}} \quad (9)$$

where

$I.P. (i,m)$  = interaction potential between growth point  $i$  and  $m$  regional markets,

$P_i$  = population of growth point  $i$ ,

$P_j$  = population of growth point  $j$ .

$t_{ij}$  = minimum path travel time between growth point  $i$  and regional market  $j$ .

For use in the calibration of the path model a modified "interaction potential" is obtained by normalizing the interaction potential by the population of city under consideration. The modified equation 10 still is a good measure of potential for interaction of city  $i$  in respect to all other growth points in the study areas but does not multiply the resultant index with the population of the city for which the index is being calculated. The basic reason for this omission is that city population is being used as a separate variable in the model and direct use of equation constitutes a double count of population.

$$I.P.(i,m)_{\text{modified}} = \sum_{j=1}^m \frac{P_j}{t_{ij}} \quad (10)$$

The calculation of modified interaction potential for sample cities required population data for all the growth centers in the region as well as the minimum travel time between the sample cities



and all other growth centers in the region. The region used for purpose of calculation of interaction potential is larger than the primary study area which is limited to the state of Indiana and the population of the growth centers often include more than population within one city limit.

#### Limits of the Study Region for Calculation of Interaction Potential

The study region for calculation of interaction potential as a proxy for inter-city accessibility is larger than the State of Indiana (the primary study area) because Indiana cities have continuous economic and cultural exchange with growth centers outside the State of Indiana.

As the distance between an Indiana city and a foreign city increases, the transportation system outside Indiana plays a more important role in the potential for interaction and the differences between the locational advantage of an Indiana city with respect to a highway system inside Indiana becomes less important in the calculation of interaction potential. Therefore, the study region for calculation of modified interaction potential was limited to the State of Indiana and parts of the neighboring states.

Two different criteria were used to choose the portions of the study area outside the State of Indiana: first, cities in the neighboring state which were located in the two counties adjoining Indiana were included in the study area, second, other cities with population of 100,000 or more and within 100 miles of Indiana borders were included.



The last criterion was time dependent and results in slightly different study areas for the study periods of 1940, 1950, 1960 and 1970. As an example, the study area for 1940 is shown in Figure 15. which is a reproduction of the coded 1940 highway network. The variable portion of the study areas for other study periods are not graphically presented but the cities included in each of the four study periods are shown in Tables A2 to A5 of Appendix A.

#### Minimum Travel Time

Travel times ( $T_{ij}$ ) between each of the sample cities and all other cities in the study area was required for calculation of interaction potentials. The minimum travel time between two cities depends on highway distance between the two and highway travel speeds (running speeds).

Once highway speeds and distances were known then the highway system was coded and a "minimum tree algorithm" applied to calculate the minimum travel time between different locations within the network. In this project, the historical highway travel speeds were not readily available and, as explained in the following sections, were estimated by assuming that the ratio of running speed and spot speed for a particular highway type was a constant from 1940 to 1960.

#### Highway Speeds

Highway speeds depend on a complex set of variables but generally were increasing from 1940 to 1970 and at a given time were higher on highways with higher design standards.



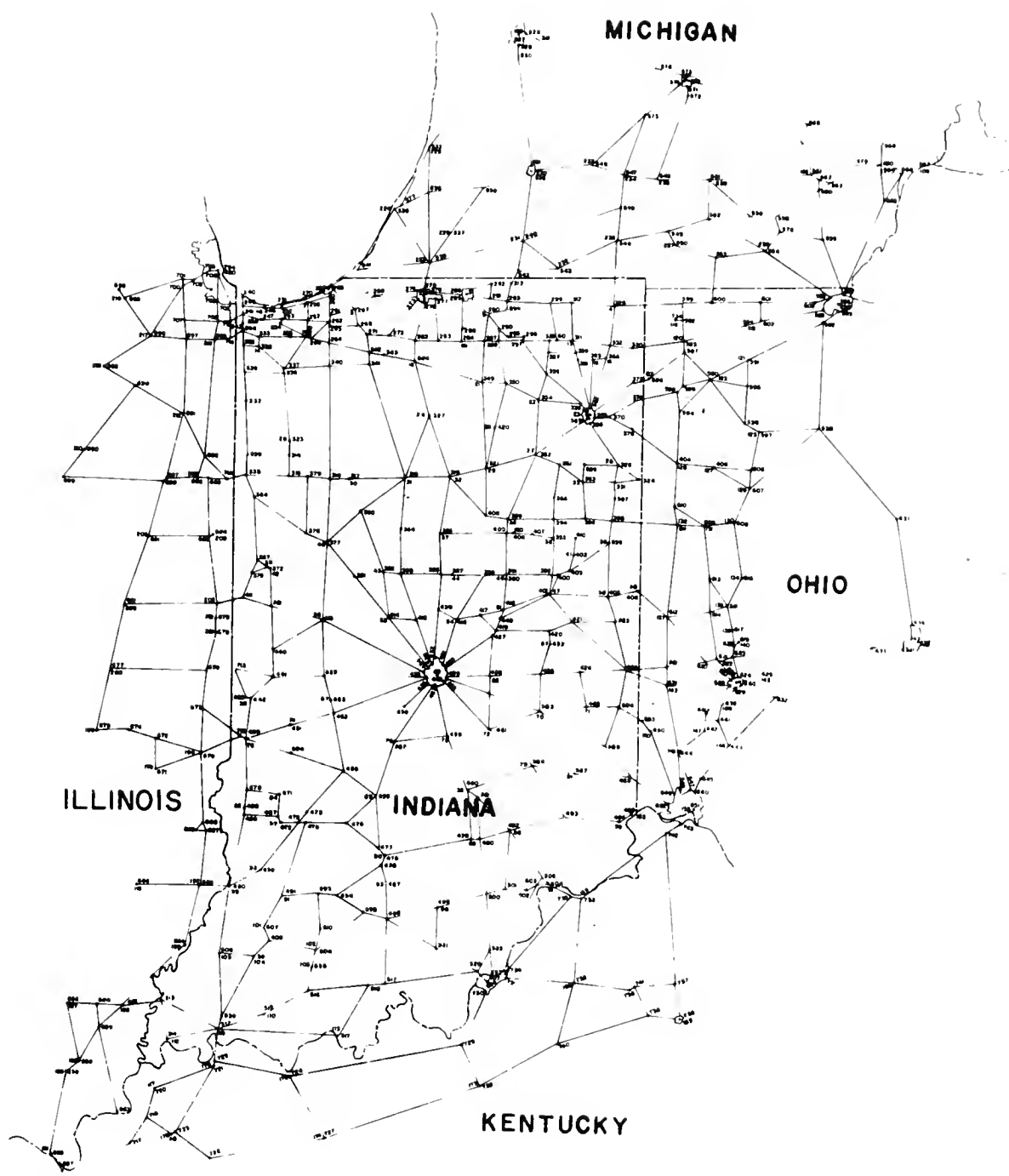


FIGURE 15. CODED REGIONAL HIGHWAY NETWORK USED FOR CALCULATION OF INTERACTION POTENTIAL-1940





Information sources available for determination of highway speeds for the four study periods were traffic speed reports prepared by the Joint Highway Research Project (10,11,12), a research report by Mr. Frost (13) which gives information about 1940 highway speeds and Vodrazka's Ph.D. thesis (14) which serves as a guide for choice of appropriate highway travel speed as a function of published spot speeds.

Vodrazka studied statewide intercity travel in Indiana for the purpose of determining the future need for upgrading of certain Indiana highways to freeway standard. In his 1966 study Vodrazka assigns travel speeds to different classes of Indiana highways as shown in Table 21.

Eighty-five percentile speeds are available for 1960 and 1970 (10,11). Speed changes between 1960 and 1970 as well as linear regression lines representing spot speed trends for different types of highways were previously calculated (10), and are shown in Figure 16.

Lines with the slopes of 1960-1970 spot speed trend line were passed thru corresponding travel speeds used by Vodrazka to yield an estimate of a reasonable travel speed for 1960 and 1970. Rural travel speeds for two-lane, four-lane and Interstate highways were thus obtained from Figure 16 and are shown in Table 22.

Travel speeds for 1950 were estimated by adjustment of June 1950 spot speeds. The ratio of mean spot speed and estimated travel speed for 1960 is shown in column 4 of Table 23. It was assumed that these ratios were approximately valid for 1950 and were multiplied with the 1950 mean spot speeds to result in 1950 estimated rural travel speeds for two lane and four lane highways which are shown in column six of Table 23.



Table 21. Assumed Highway Travel Speeds - 1966

Highway Type	Travel Speeds (mph)
Interstate	60
4-Lane Rural	50
2-Lane Rural	45
4-Lane Urban	30
2-Lane Urban	15

Source: Vodrazka, Walter Charles, "Subclassification of the State Highway System of Indiana Based on Synthesis of Intercity Travel", Ph.D. Thesis, Purdue University, W. Lafayette, Indiana, June 1968.



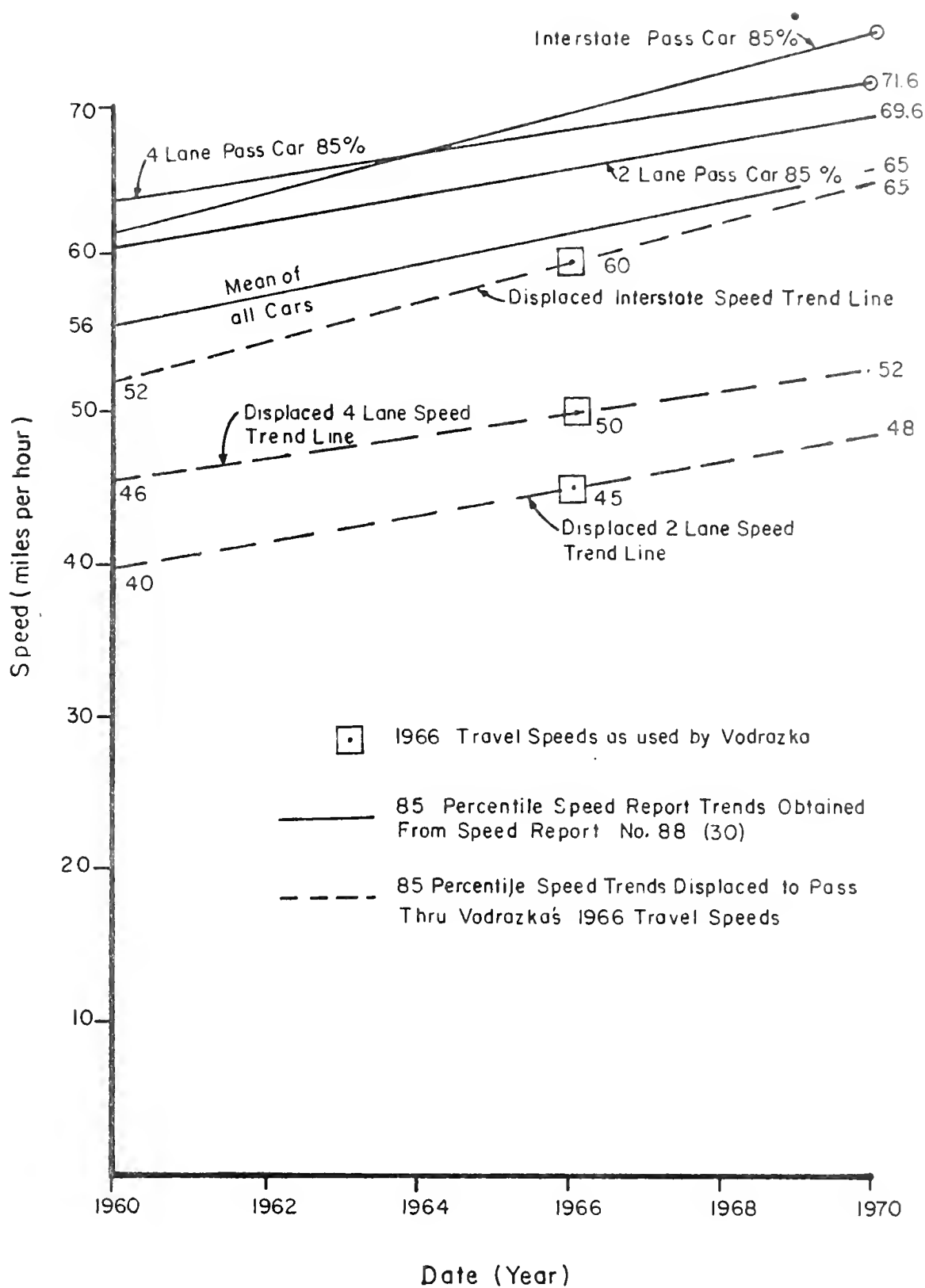


Figure 16. Rural Travel Speed 1960-1970



Table 22. Estimated Rural Highway Travel Speeds - 1960 and 1970

Highway Type Column 1	1960		1970	
	Measured 85% Spot Speed (mph) Column 2*	Estimated Running Speed (mph) Column 3*	Measured 85% Spot Speed (mph) Column 4*	Estimated Travel Speed (mph) Column 5*
Interstate	62	52	75	65
4-Lane	64	46	72	52
2-Lane	60	40	70	48

\*Source Figure 16.

Table 23. Estimation of Average Travel Speeds for Rural Highways in Indiana - 1950

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
Highway Type	Mean Spot Speed for All Autos in 1960(1) (mph)	Assumed Travel Speed for 1960(1) (mph)	Ratio of Travel Speed to Spot Speed for 1960: (Col.3/Col.2)	Mean Spot Speed 1950 (2) (mph)	Estimated Travel Speed 1950: (Col.4xCol.5)
2-Lane	56	40	0.72	52	38
4-Lane		46	0.82	56	46

Notes: (1) Values obtained from Figure 16.  
 (2) Source: Traffic Speed Report 44 (Reference 12).





Frost in a study of "Indiana war time travel speeds" gives the overall 50 percentile spot speed measured in February 1942 as 49.56 miles (13). Assuming a linear speed trend between 1940 and 1960 the overall 50 percentile spot speed was adjusted to 46 mph for 1940. A further assumption that the ratio of 50 percentile spot speed to running speed remains the same as that of 1960 (shown in Table 22) allows the estimation of a 33 mph running speed ( $46 \times 0.72$ ) for 1940 two-lane rural highways.

Historical highway speeds in urban areas were not available but it was necessary to assume urban highway speed to enable "loading" the network. Assumed urban highways travel speeds for the four study periods are shown in Table 24. It is believed that these assumptions are reasonable. Minimum travel times are not very sensitive to urban travel speeds because the magnitude of intercity travel over urban highways are usually less than one mile.

#### Coding the Highway Network

Highway maps of Indiana (15,16,17,18), Illinois (19,20,21,22), Kentucky (23,24,25,26), Michigan(27,28,29,30) and Ohio (31,32,33,34) were used to code the regional highway network for the four study periods of 1940, 1950, 1960 and 1970. Highways on these maps were classified as to number of lanes and type of access. Therefore, highway travel speeds obtained in the last section were used to assign travel speeds to the highway links. Highway distances were usually read from the map or actual highway mileage were scaled. The scale of the maps used ranged from 1 in. = 7.9 miles to 1 in. = 16 miles which



Table 24. Assumed Urban Highway Speed for Coding Inter-City Highway Network\*

Highway Type	1940 Speed (mph)	1950 Speed (mph)	1960 Speed (mph)	1970 Speed (mph)
Two-Lane	15	15	15	20
4-Lane	N.A.	N.A.	27	32
Urban Freeway	N.A.	N.A.	40	40

\*Urban highways are used only for loading the inter-city network.

N.A. - not applicable. The highway maps available do not distinguish between two-lane highways and multi-lane highways in the urban areas.



allowed scaling of highway distances to a tenth of a mile. A few highway links in the 1940 network were not paved with asphalt or concrete and their travel speed was penalized from 33 miles per hour to 28 miles per hour.

Growth centers with greater than 2,500 population were coded as a centroid. The population of the growth center included that of the central city as well as that of the fringe areas. Each of the urbanized areas and standard statistical areas in the region were considered as one growth center. For other places, the population of villages and towns and cities as close as three miles to each other were aggregated to a single centroid. The number of centroids as well as the number of nodes and links for the network of each of the four study periods is shown in Table 25.

Centroids were connected to the network nodes by dummy links. The dummy links were used to load the network and were assigned a distance to account for the distance from the homes to main city arterials. For urbanized areas these distances were assumed to be one-fourth the city's average dimensions. For smaller centroids, the following impedance distances which were obtained from Vodrazka (35) were assigned to the dummy links. Also some highway links contained one or more towns with less than 2,500 populations. Passing thru these small population centers cause travel delay which was accounted for by increasing the actual highway mileage by the same impedance distances.



Table 25. Size of Coded Regional Highway Network - 1940 to 1970

Network Element	Number* of each element of the network in the year			
	1940	1950	1960	1970
Centroid	175	194	218	234
Nodes	744	903	856	1354
Links	2010	2576	2386	4212

\*The numbers given are approximate. It was necessary to keep the same centroid numbers for each decade. Therefore originally centroids for 1970 were numbered from 1 to 239. When a decrease in population caused a 1970 centroid not to be considered a centroid in 1960 network its number was kept and it was considered as a dummy centroid. There are 4 dummy centroids in 1970, 11 dummy centroids in 1960, 70 dummy centroids in 1950, and 84 dummy centroids in 1940 highway network. Each dummy centroid has two associated dummy links. Additionally there are some other extra nodes.





<u>City Size</u>	<u>Impedance Distance</u>
Over 5,000	1.0 mile
1,000-5,000	0.5 mile
100 -1,000	0.2 mile

It was assumed that all links were two way links and distance as well as speed on both directions was the same. As a result of this assumption only one way links were coded and a small computer program was written to read the four pieces of information for each link (Node 1, distance, speed, Node 2), calculate the travel time and re-write the complete information for both directions of travel. Next, a sort routine was used to rearrange the links in the increasing order of Node 1 and then on the increasing order of Node 2. The computer program TNET-5 (36) was used to check the network to assure that all the links were two way, ordered correctly, and there were no missing nodes or centroids. Once a correct network was available then it was used to calculate the minimum tree paths between the centroids.

The plans of coded 1940, 1950, 1960 and 1970 highway network as well as the listings of corresponding coded highway network are available in the Purdue University Joint Highway Research Program. As an example, and to show the limits of the study area, the 1940 coded highway network is shown in Figure 15.

#### Algorithm Used to Calculate Minimum Travel Time Between Centroids

The coded highway networks were used to calculate the minimum travel paths between the centroids. The algorithm used was that of



the British Road Laboratory, as used in the Tnet-10 computer program developed by the University of California at Berkeley (37). The dimensions of this program were modified to enable the solution of minimum path trees for relatively large networks analyzed in this project.

The output of this program for each study period was the minimum path trees from each centroid to all other centroids and a minimum time matrix between the centroids. These outputs are not presented here because of the space limitation, but are stored on magnetic tape for future reference. The minimum time matrix for each period then was used in calculation of the interaction potential for the respective time period.

#### Calculation of Interaction Potentials

Once the minimum travel times were calculated, interaction potential in accordance with equation 10 was calculated. The other data requirement for this equation was the population of the centroid which was obtained from previously referenced census publications (3,4, 5,6).

Program "potential", shown in Table A1 of Appendix A, was written to compute interaction potentials for all the centroids in the study area. Interaction potentials for the sample cities are shown in column 4 of Tables 26 to 29. Column 3 of these tables show the population of the centroid used in calculation of the interaction potentials. These populations were larger than the population of the city proper used in the path model and given in Tables 37 to 40.



Table 26. Interaction Potential for 1940 Sample Cities

Centroid No.	Name of Centroid	Population of Centroid	Interaction Potential
11	VALPARAISO	8736	6.263703E+04
13	KENDALLVILLE	5431	4.454491E+04
21	WARSAW, WINONA LAKE	7121	4.582090E+04
27	HUNTINGTON	13903	4.478665E+04
37	KOKOMO	33795	4.310595E+04
38	HARTFORD CITY	6946	4.253133E+04
40	LAFAYETTE, WEST LAFAYETTE	35068	4.402127E+04
42	ATTICA, WILLIAMSPORT	4982	4.145354E+04
43	FRANKFORT	13706	4.255829E+04
47	MUNCIE METROPOLITAN AREA	49720	4.360118E+04
58	RICHMOND, CENTERVILLE, BOSTON	36491	4.226820E+04
67	GREENCASTLE	4872	3.812224E+04
71	CONNERSVILLE	12898	3.917343E+04
72	SHELBYVILLE CITY	10791	4.156110E+04
79	GREENSBURG	6065	3.891172E+04
81	BATESVILLE	3065	3.754302E+04
83	BLOOMINGTON	20870	3.555374E+04
87	LINTON	6263	3.168982E+04
90	BEDFORD , OOLITIC	13700	3.301499E+04
95	VINCENNES	18228	3.058113E+04
102	JASPER	5041	2.788404E+04



Table 27. Interaction Potential for 1950 Sample Cities

Centroid No.	Name of Centroid	Population of Centroid	Interaction Potential
24	ROCHESTER	4673	5.006352E+04
30	MONTICELLO	3467	4.882128E+04
31	LOGANSPOUT CITY	21031	4.971792E+04
37	KOKOMO	38672	4.993729E+04
38	HARTFORD	7253	5.061418E+04
39	PORTLAND	7064	4.931728E+04
43	FRANKFORT	15028	4.954254E+04
44	TIPTON	5633	5.018552E+04
46	ALEXANDRIA, ORESTES	5629	5.237169E+04
51	ANDERSON, CHESTERFIELD, EDGEW	48702	5.274507E+04
57	NEW CASTLE	18271	4.884838E+04
58	RICHMOND, CENTERVILLE, BOSTON	41182	5.303423E+04
67	GREENCASTLE	6888	4.417765E+04
72	SHELBYVILLE CITY	11734	5.117967E+04
73	FRANKLIN	7316	5.003118E+04
83	BLOOMINGTON	28163	4.071288E+04
86	AURORA	4780	4.824035E+04
90	BEDFORD , OOLITIC	13687	3.825397E+04
92	BICKNELL	4572	3.506808E+04
95	VINCENNES	18831	3.577351E+04
101	PETERSBURG CITY	3035	3.382525E+04
102	JASPER	5215	3.247754E+04
115	TELL CITY	5735	2.964402E+04
151	NAPANEE	3393	5.128927E+04





Table 28. Interaction Potential for 1960 Sample Cities

Centroid No.	Name of Centroid	Population of Centroid	Interaction Potential
1	MICHIGAN CITY, LONG BEACH+TWO	41151	9.403323E+04
9	GOSHEN	13718	8.419606E+04
18	PLYMOUTH	7558	8.832902E+04
31	LOGANSPOUT CITY	21106	7.710206E+04
34	BERNE CITY	2644	6.932143E+04
40	LAFFAYETTE, WEST LAFFAYETTE	55010	7.053453E+04
44	TIPTON	5604	7.767662E+04
45	ELWOOD	11793	7.527371E+04
50	COVINGTON	2759	6.589713E+04
51	ANDERSON, CHESTERFIELD. EDGEW	53768	7.112907E+04
51	ANDERSON, CHESTERFIELD. EDGEW	53768	7.112907E+04
53	LEBANON, ULEN	9904	7.253246E+04
54	NOBLESVILLE	7664	8.131942E+04
58	RICHMOND, CENTERVILLE, BOSTON	46767	7.664792E+04
72	SHELBYVILLE CITY	14317	7.432396E+04
76	MARTINSVILLE	7525	6.957592E+04
78	EDINBURG	3664	6.599531E+04
82	COLUMBUS	20778	6.333144E+04
83	BLOOMINGTON	31357	6.321825E+04
91	MADISON, HANOVER	11658	5.797822E+04
92	BICKHELL	3878	5.231502E+04
96	WASHINGTON	10846	5.261159E+04
101	PETESBURG CITY	2939	4.999482E+04
110	BOONVILLE	4801	4.625213E+04
162	LAWRENCEBURG, GREENDALE, HARDIN	8083	5.302948E+04
164	SCOTTSBURG	3810	6.331375E+04



Table 29. Interaction Potential for 1970 Sample Cities

Centroid No.	Name of Centroid	Population of Centroid	Interaction Potential
1	MICHIGAN CITY, LONG BEACH+TWO	45813	1.323261E+05
1	MICHIGAN CITY, LONG BEACH+TWO	45813	1.323261E+05
3	ELKHART CITY	43152	1.462362E+05
9	GOSHEN	17171	1.457876E+05
19	KNOX	3519	1.200493E+05
35	FOWLER	2643	1.093973E+05
38	HARTFORD	8207	1.159268E+05
39	PORTLAND	7115	1.182500E+05
50	COVINGTON	2641	9.975936E+04
52	CRAWFORDSVILLE	13842	1.053151E+05
56	UPLAND	3202	1.205036E+05
57	NEW CASTLE	21215	1.132631E+05
58	RICHMOND, CENTERVILLE, BOSTON	46589	1.120502E+05
71	CONNERSVILLE	17604	1.094747E+05
82	COLUMBUS	27141	1.013708E+05
83	BLOOMINGTON	42890	8.947667E+04
91	MADISON, HANOVER	16099	8.913003E+04
92	BICKNELL	3717	7.267331E+04
96	WASHINGTON	11358	7.646498E+04
102	JASPER	8641	7.262491E+04
109	FORTBRANCH	2535	7.658658E+04
110	BOONVILLE	5736	7.499043E+04
111	BLOOMFIELD	2565	8.066977E+04
157	DELPHI	2582	1.084019E+05



The project required only interaction potentials for the sample cities, but the minimum time matrix and the population information for all the centroids in the region were required for calculation of the interaction potentials for these sample cities; therefore, it was decided to calculate the interaction potential for all the centroids in the region. Availability of interaction potential for all Indiana cities would give a better feel for this variable and in case of future extension of the study to other cities the interaction potentials would be readily available. The interaction potential for all centroids from 1940 to 1970 are presented in Tables A2 to A5 of Appendix A.

#### Intra-City Accessibility (X3)

It was desired to obtain an index of overall accessibility for the transportation system for each sample city. It was noted that the primary mode of travel in the small Indiana cities being studied was by highway using the personal auto. The accessibility index (51)

$$acc_i = \sum_{j=1}^n a_j f_{ij} \quad (12)$$

where

$acc_i$  = accessibility index

$a_j$  = number of attraction

$f_{ij}$  = travel time factor between zones i and j.

used in urban transportation planning was a satisfactory measure for this purpose, but unfortunately the level of aggregation of the experiment did not allow its deployment for calculation of intra-city



accessibility. The scope of the experiment allows only the collection of data on the level of the city as a whole and it was not possible to collect zonal attractions or minimum time matrix between the city zones. Therefore, another measure of overall intra-city accessibility was needed.

Consideration was given to use of lane-miles of highways within the city adjusted according to their functional classification. The use of this measure for intra-city accessibility was also rejected because of the difficulty of adjusting the actual highway lane miles in a uniform (and repeatable) manner for their functional classification. Finally acres of land deployed for streets and highways within each city were used as proxy for intra-city accessibility. Attempts were made to estimate total land in the highway right of way for this purpose. It is often possible to estimate the limit of right of way from such signs as sidewalks, yard lawns, fences, drainage ditches, and cultivation of abutting lands and knowledge of approximate widths of right of way.

This operational definition of intra-city accessibility places land resource used for highways in the same level as other competing land uses as defined in the next paragraph. Therefore the numerical values for this data were collected with the data for other land uses as discussed in the next section.

#### Land Use Variables (X4, X8, X9, X10, X11)

Industrial land use (X4), commercial land use (X8), vacant land use (X9), institutional land use (X10), and residential land use (X11)





were operationally defined for purpose of calibration of the path model in acres of land used for each purpose as measured from aerial photographs.

Prior to the land use measurement Table 30 was prepared to classify land uses under nine categories to assist in the uniform land use classifications in the airphoto interpretation phase of this experiment. Nine land use categories of Table 30 were chosen as to be compatible with the detail available on the aerial photographs with a scale ratio of 1:20,000. These categories were also compatible with land use classification usually used by city planners and were useful in the calibration of the path model. Actually the land categories shown in Table 30 are more detailed than required for calibration of the path model. Lands categorized as other transportation, agriculture, and waste and forest were not used in the calibration of the path models, the reason for more detailed classification in data collection than the one used in the path model was to have some flexibility in case of later needs for revision of the path model. Also with the land use measurement method used all the city's land had to be accounted for.

#### Method of Measurement of Land Use Data

All land uses were obtained from vertical aerial photographs with a scale of 1:20,000 (1 inch = 1666.66 feet) published by the United States Department of Agriculture (38,39). This scale is sufficient for study of land use changes in urban places (40). Most of the required 1940 photos were available in the Joint Highway Research Project Laboratory,



Table 30. Classification Used in Land Use Measurement

Land Uses Were Recorded in These Categories	Land Use Sub-Classification
1. Residential Land Use	Single Family Residential Multi-Family Residential Mobile Homes
2. Commercial Land Use	Retail Establishment Office Buildings Shopping Centers Wholesale
3. Industrial Land Use	Manufacturing Mines and Quarries
4. Institutional Land Use	Institutional Buildings (schools, government buildings) Institutional Open Space (ballparks, parks, country clubs) Cemeteries
5. Streets and Highways	Arterials Highways Collector Streets Local Streets
6. Other Transportation	Railroad Airports Water Transportation Utilities (transmission line-power plants-sanitation)
7. Vacant and Unused	Land within city which is not being used, but from form and slope it seems easy to build on
8. Agriculture	Agriculture (includes land used for barns and other farm buildings)
9. Water and Forest	River, Lake, Filled Quarries Woody Areas Other areas with rough terrain un- suitable for building



but many of the photos for 1950, 1960, and 1970 were borrowed from Indiana State Highway Commission, and U.S. Geological Survey office at Bloomington, Indiana.

Land uses measured from the aerial photographs were to be associated with their corresponding socio-economic data obtained from continental census data. Therefore it became necessary to draw on the aerial photos the corresponding city limits used by the census enumerators. The census enumerators used the regular city limits for census boundaries, but for cities under 50,000 population these city limits were recorded in unpublished "place maps" which for the first ten years are made available to the public thru special order from the Census Bureau. Thereafter the "place maps" may be purchased from the National Archives and Record Service. Extensive use of "place maps" were made in this project because the city limits were changing, and usually it was not possible to obtain from the cities a map for the specific dates desired. Most cities, evidently, were updating the city limit on the same original and did not try to keep a record of city limit changes. The required 1970 place maps were borrowed from the Geography Department Library at Indiana University. For other dates in lieu of a "place map", which usually was purchased for the project, when possible the city limits shown on the United States Geological Survey Maps (U.S.G.S.) were used.

The city maps obtained from different sources were used to draw the city limits on the corresponding aerial photographs. The city limits used were usually very close to 1940, 1950, 1960, or 1970, but the aerial photographs were often as much as two to three years



different than the census dates. This variation in the dates of aerial photographs were assumed acceptable, because land uses usually do not change rapidly.

The vertical aerial photographs used were most accurate in the center and become distorted toward the edges. But fortunately, because of the 50 percent or more overlap of the stereo pairs, mostly the center of the photographs were used for land use determination through statistical point sampling and airphoto interpretation.

#### Statistical Point Sampling

Statistical point sampling was used in lieu of delimiting individual land uses on the photos and subsequent area measurement. The statistical point sampling has been used by the Chicago Area Transportation Area (CATS) for suburban land use inventory (41) and others (42). In the statistical point sampling a previously designed pattern of dots on a transparency is overlaid on the aerial photos of the city and land use under each dot is recorded. The method estimates proportion of population of points in a given area which fall within each land use. The proportions are converted to acreage of land use by determining total acreage of the area being inventoried. The standard error of the sample proportion depends on the density of points, and was determined from equation 11 used by CATS (41).

$$\sigma_p = \sqrt{\frac{pq}{N}} \quad (11)$$

where

$\sigma_p$  = standard error of sample proportion

P = the sample proportion of point in the land use for which variance is being measured





$q$  = the proportion of other use

$N$  = the total number of points falling in a city.

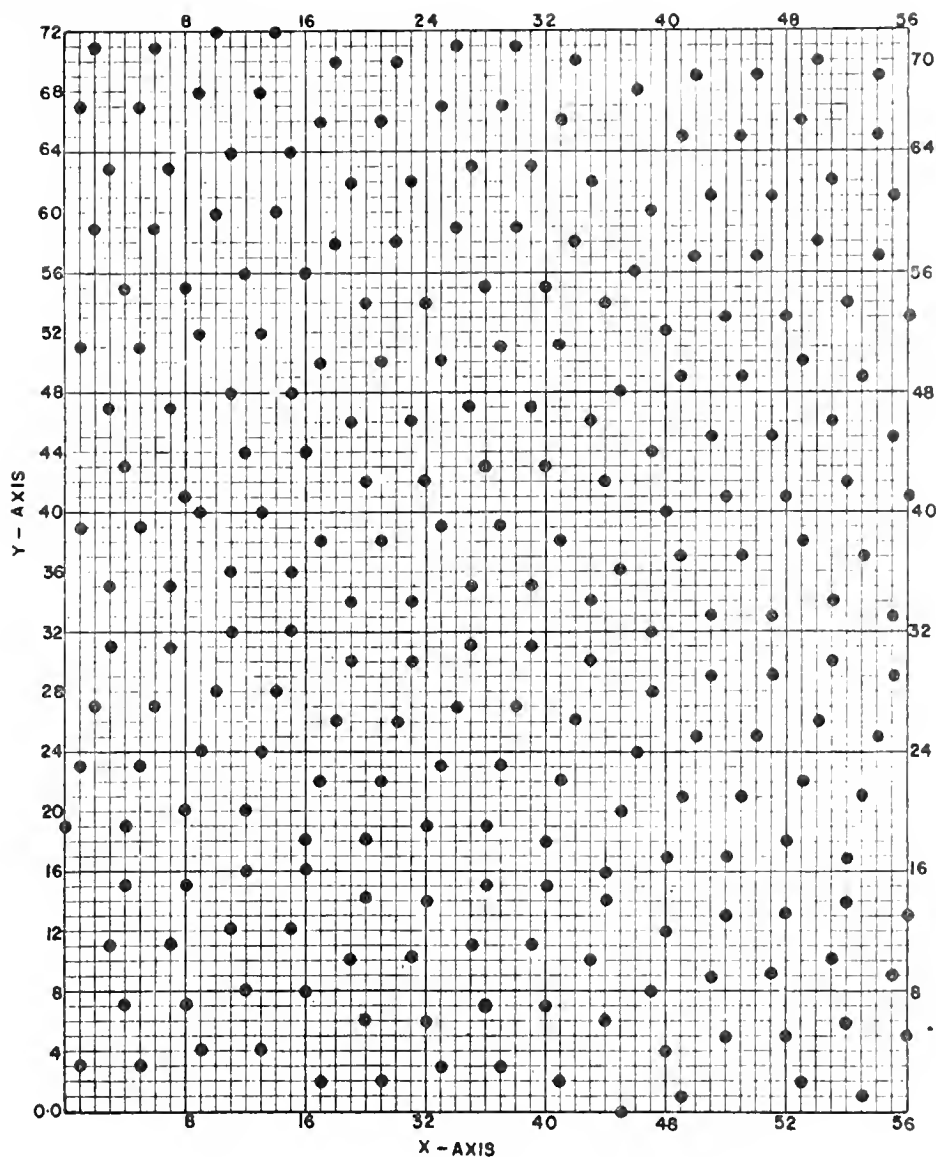
The above equation assumes randomness of points. Actually the points are not chosen randomly, but the points were designed to yield a "stratified systematic unaligned sample", and according to CATS this sampling scheme causes the above equation to be conservative. The details of construction of a stratified systematic unaligned sample overlay was explained in the Agricultural Handbook No. 237 (43). This is an efficient sampling procedure because the resultant stratified systematic sampling will provide unbiased estimates due to lack of periodicities.

The dot pattern overlay designed for the current project is reproduced in Figure 17 which shows four points were sampled for every one square inch of aerial photographs. Given the scale of aerial photos of 1:20,000 this sampling density was slightly over 40 dots (samples) per square mile. This sampling density compares with 44 points per square mile used in the CATS suburban land use survey.

#### Airphoto Interpretation

The decision as to the type of land use under a particular dot was made by the use of the technique of "airphoto interpretation". This method was explained in several references (45, 46) and used for urban land use collection by others (47, 48, 49). Airphoto interpretation of a particular land use was made by use of a stereoscope and a stereo pair of photos (photos which overlap in the direction of





- Notes: 1) To show the dot pattern clearly the individual dots are shown larger than the ones used in overlay.
- 2) The figure shown is on a reduced scale. On the actual overlay used each eight units of X or Y axis was equal to one inch.

Figure 17. Dot Patterns Designed for Statistical Sampling



flight) which allows three dimensional observation of the terrain and the buildings. Most of the land use under the dots were interpreted with the aid of a mirror stereoscope with an attached telescope which magnified the objects in the photograph three times.

#### Test of Statistical Point Sampling Method

Prior to the land use data collection by point sampling technique the density of sampling, as well as the skill of the investigator in airphoto interpretation were to be tested. This was accomplished by comparing 1969 land uses for Richmond, Indiana obtained from aerial photographs by statistical point sampling method with 1969 land uses obtained by the standard land use inventory method. The values for inventoried land use were obtained from the planning commission of city of Richmond.

The city limits of Richmond was outlined directly on 12 aerial photographs with red china marking pencil. (China pencils are erasable without damaging the photos). Additional green lines assigned exclusive sections of the city to specific photographs in such a manner as to not use distorted edges of the photographs. Dot pattern transparency was overlaid on each photograph and a record made of the relative position of the overlay and photograph to establish a reference for X and Y coordinates of the overlay. Next the land use immediately under the dot was classified to one of the land use categories defined in Table 30. A record was made of the land use under the dot as well as the X and Y coordinates of the dot. The reason for recording the coordinates of each dot was to enable duplication of the measurement process.



The results of 560 statistical point samplings of land uses in the city of Richmond were the number of dots which had fallen on top of nine different land uses. Given the fact that the density of the dot patterns used were four dots per one square inch of the aerial photograph and the fact that the central area of the airphotos have an approximate scale of one inch to 1667 feet then every dot represents 15.94 ( $1667 \times 1667 \div 4 \times 43560$ ), acres of land. Therefore, the number of dots in a given category of land use were multiplied by 15.94 to convert the measured dots to its equivalent statistical area in acres.

The land use inventory for Richmond and statistical point sampling method were compared in Table 31. Land use types as listed in column two of this table were obtained from the Richmond Planning Commission. Industrial land use and land use for other transportation in Table 30 were aggregated for incorporation in Table 31; as were vacant, agriculture, and water and forest in Table 30 were aggregated for use in Table 31 under classification of undeveloped land. Column 4 of Table 31 shows the standard error of each measurement calculated by equation 11 as exemplified in the bottom of the Table 31 where the standard error for residential land use was calculated. Column 6 shows the 95 percentile confidence interval which were obtained based on the assumption of normality of dot patterns. Although the standard error used in construction of this confidence interval was obtained from the point sampling method, the inventoried land use values were used as best estimates of the means of each land use. Column 7 of Table 31 shows the percentage of each land use as measured from sampling of aerial photographs. Comparison of Column 6 and





Table 31. Comparison of Statistically Point Sampling with Inventory Method of Land Use Measurement

Land Use Type Column 1	Measurement Ob- tained by Stan- dard Land Use Inventory Method (Acre) Column 2	Measurement Ob- tained by Statis- tical Point Sampling of Aerial Photos (Acre) Column 3	Standard Error of Point Sampling* Column 4	Inventoried Land Use as Percent of Total Land Use (%) Column 5	95 Percent Confidence Interval for Each Land Use (%)** Column 6	Statistically Sampled Land as Percent of Total Land Use (%) Column 7	Do Sampled Land Land Uses Fall in the Confidence Interval for the Inventoried Land Use Column 8
Residential	2538.6	2829	1.9	28.99	25.2 - 32.8	32.3	Yes
Streets	1100.	1157	1.4	12.57	9.8 - 15.4	13.2	Yes
Industrial + Railroad + Utility	733.26	828	1.2	8.38	6.0 - 10.8	9.5	Yes
Commercial	555	672	1.0	6.34	4.3 - 8.3	7.7	Yes
Institutional	1034.64	797	1.4	11.82	9.8 - 14.6	9.1	Yes
Undeveloped	2792.5	2470	2.0	31.90	27.9 - 35.9	28.2	Yes
TOTAL	8754						

\*Standard error is calculated by  $\sigma = \sqrt{pq/N}$  (equation 11) as an example standard error for residential land use is calculated:

$\sigma_R$  = standard error for residential land use

p = proportion of residential land use = 0.2899

q = proportion of other land uses = 1 - 0.2899

N = total number of points sampled in the city of Richmond = 560 points

$\sigma_R = \sqrt{(0.2899)(1-0.2899)/560} = 0.019$  or 1.9%.

\*\*Confidence Interval = Inventoried Land Use  $\pm Z_{\alpha} \sigma$  (Standard error).



Column 7 shows that for all land use categories the statistically measured land use was within the 95 percentile confidence limits of corresponding inventoried land use.

#### Repeatability of Statistical Point Sampling

To check on repeatability of statistical point sampling method, six of the twelve photos covering the city of Richmond were randomly chosen and the whole statistical point sampling procedure with the exception of drawing the city limit lines on the photos was repeated. Results of the second partial land use determination are compared in Table 32 with the land use measured on this same six photos on the first statistical point sampling of Richmond. Comparison of columns six and seven of this table shows that the land use measured in the second run was within the 95 percent confidence interval of the measurement of the first run. Therefore it was assumed that statistical point sampling may be used as an efficient means of obtaining land uses for calibration of the transportation-city development path model.

#### Collection of Land Use Data

The land use data were collected for all but two of the sample cities by the statistical point sampling method described for Richmond. The two cities for which land use data were not collected are the cities of Dunkirk (1940) for which city limits could not be located (National Archives could not locate its place map) and Rockville (1960) for which aerial photos were not available. These two cities were not replaced and it was assumed that their omission was not affecting the experiment significantly.



Table 32. Testing Repeatability of Land Use Determination by Statistical Point Sampling of Aerial Photographs

Land Use Type Column 1	Land Use Measured from Six of the Twelve Aerial Photos of Richmond, Indiana		Percentage of Individual Land Uses Measured from the Six Randomly Chosen Photos Based on Totals of First Run		Standard Error for Second Run* (Percent) Column 6	95 Percent Confidence Interval for the First Run (Percent) Column 7	Do Land Use Measurements in the Second Run Fall in the Confidence Interval for the First Run Column 8
	First Run (Acre) Column 2	Second Run (Acre) Column 3	First Run Column 4	Second Run Column 5			
Residential	1594	1674	25.0	26.3	2.2	20.7 - 29.3	Yes
Commercial	358	478	5.6	7.5	1.17	3.3 - 7.9	Yes
Industrial	686	558	10.8	8.8	1.58	7.7 - 13.9	Yes
Institutional	661	606	10.4	9.5	1.56	7.3 - 13.5	Yes
Streets	847	749	13.3	11.8	1.73	9.9 - 16.7	Yes
Other Transportation	288	239	4.5	3.8	1.06	2.4 - 6.6	Yes
Undeveloped Land	1931	1817	30.3	28.6	2.34	25.7 - 34.9	Yes
TOTAL	6365						



Prior to the actual data collection attention was paid to the order in which the sample cities were to be measured. The ordering of sample cities was important to avoid the error of "instrumentation". Error of "instrumentation" is the error which causes bias in the result of experiment due to systematic data measurements. It is not easy to always identify the sources of the error of "instrumentation", therefore it is useful to randomize the order in which the data for sample cities is measured.

Airphotos for 1950, 1960, and 1970 sample cities were borrowed from the Indiana State Highway Commission and U.S. Geological Survey. It was not desirable to request to borrow all the photos for a long time. Therefore, as a compromise, cities in every study period were divided randomly into three batches and the airphotos for these batches were interpreted continuously, but with internal re-randomization.

Tables 33 to 36 show land use measurements for the sample cities in 1940, 1950, 1960 and 1970 respectively. Column 1 of these tables shows the last two digits of the year of the aerial photographs (as well as its month in decimal) used in the land use measurements. Columns 3 to 11 shows the acreage of different land uses measured from aerial photos for sample cities. The classification shown in these columns was the same as the one explained in Table 30.

#### Census Data (X5, X6, X7)

Total city population (X7), total employment (X5) and median family income (X6) for sample cities were obtained from census





Table 33. Land Use Measurements for 1940 Sample Cities in Acres

City Name	Date of Aerphotos	Intra-City Highways	Industrial	Commercial	Institutional	Residential	Other Transportation	Agriculture	Water and Forest	Vacant
VALPARAISO	38.8	319	32	15	64	415	64	144	32	446
KENDALLVILLE	38.8	175	96	16	48	367	64	239	64	159
WARSAW	38.9	255	32	16	32	446	48	80	255	239
HUNTINGTON	39.7	494	80	48	112	765	159	96	128	367
KOKOMO	38.6	606	191	128	96	1371	*	*	175	303
HARTFORD CITY	41.6	207	*	64	32	415	32	96	32	32
LAFAYETTE	39.3	574	80	80	303	1451	112	271	64	175
ATTICA	39.5	144	16	16	16	207	*	175	80	48
FRANKFORT	39.4	606	32	64	128	510	48	159	48	144
MUNCIE	41.6	1259	399	16	638	1929	144	255	255	733
RICHMOND	40.5	998	96	112	96	1212	96	446	335	383
GREENCASTLE	39.6	144	*	32	96	207	16	16	96	191
CONNERSVILLE	40.5	271	64	16	64	622	48	175	175	255
SHELBYVILLE	41.7	239	64	64	32	351	*	64	16	64
GREENSBURG	40.7	144	16	16	16	239	*	112	16	*
BATESVILLE	40.7	80	16	16	80	191	32	112	*	64
BLOOMINGTON	39.5	526	48	48	239	733	80	80	96	393
LINTON	37.5	303	16	56	48	351	16	430	64	96
BEDFORD	38.0	335	144	*	96	717	48	*	191	297
VINCENNES	40.6	383	96	80	112	415	64	144	48	128
JASPER	37.8	80	112	16	48	223	32	*	16	80

--- AT THE LEVEL OF AGGREGATION USED IN LAND USE MEASUREMENT THIS LAND USE WAS NOT OBSERVED.



Table 34. Land Use Measurements for 1950 Sample Cities in Acres

City Name	Date of Airphotos	Intra-City Highways	Industrial	Commercial	Institutional	Residential	Other Transportation	Agriculture	Water & Forest	Vacant
ROCHESTER	51.5	144	32	32	112	319	16	32	128	223
MONTECELLO	51.7	159	16	16	32	239	*	223	144	*
LOGANSPOUT	51.8	590	144	64	223	988	144	16	303	207
KOKOMO	51.0	749	303	96	223	1802	32	80	191	128
HARTFORD	50.4	175	16	16	48	542	16	128	16	48
PORTLAND	50.5	175	48	32	48	399	48	80	32	16
FRANKFORT	50.6	319	22	112	303	813	128	96	32	112
TIPTON	50.6	144	*	16	64	367	48	16	96	48
ALEXANDRIA	50.0	112	16	16	48	271	64	112	32	112
ANDERSON	50.0	909	558	175	383	2041	207	526	223	287
NEW CASTLE	50.0	367	112	32	287	861	96	96	48	48
RICHMOND	50.6	749	494	287	367	1594	80	478	271	48
GREENCASTLE	46.9	80	32	16	48	383	16	16	48	32
SHELBYVILLE	50.7	271	64	64	80	367	*	112	16	64
FRANKLIN	58.8	175	48	16	159	430	48	159	48	112
BLOOMINGTON	45.8	430	80	32	239	797	64	16	16	271
AUSORA	49.0	112	32	48	48	223	16	16	64	32
BEDFORD	46.9	367	56	16	64	686	80	16	175	144
BICKNELL	53.8	128	*	16	32	430	16	64	*	80
VINCENNES	53.8	415	112	112	207	781	32	32	32	112
PETERSBURG	49.0	96	16	16	16	191	16	16	*	16
JASPER	49.4	112	48	*	32	255	16	*	64	80
TELL CITY	53.7	128	48	48	64	271	48	32	64	48
NAPANE	51.0	144	48	32	32	255	16	207	16	16

\* - AT THE LEVEL OF AGGREGATION USED IN LAND USE MEASUREMENT THIS LAND USE WAS NOT OBSERVED.



Table 35. Land Use Measurements for 1960 Sample Cities in Acres

City Name	Date of Atrphotos	Intra-City Highways	Industrial	Commercial	Institutional	Residential	Other Transportation	Agriculture	Water & Forest	Vacant
MICHIGAN CITY	58.7	1307	446	96	1036	1722	462	1706	2933	1339
GOSHEN	57.4	430	175	32	128	702	48	223	383	191
PLYMOUTH	57.5	319	64	80	96	415	48	175	207	239
LOGANSPOET	57.7	430	80	64	415	1020	96	16	446	207
BERNE CITY	57.5	96	*	16	32	191	16	159	16	32
LAFAYETTE	57.7	287	48	32	255	558	32	239	48	48
TIPTON	56.7	144	*	*	112	335	64	32	48	32
ELWOOD	61.5	415	112	32	159	829	48	16	64	207
COVINGTON	62.0	80	*	16	48	271	32	96	112	*
ANDERSON	61.4	1052	606	80	606	2057	287	128	223	478
CHESTERFIELD	61.4	80	*	16	32	175	*	32	64	32
LEBANON	62.5	415	112	48	255	526	48	303	64	255
NOBLESVILLE	62.5	191	16	32	96	399	32	112	32	16
RICHMOND	61.4	1148	399	399	494	2009	335	1100	941	526
SHELBYVILLE	61.4	446	207	48	144	686	80	239	207	207
MARTINSVILLE	58.5	367	32	16	128	271	16	32	96	128
EDINBURG	62.7	96	48	32	32	271	16	*	48	48
COLUMBUS	60.4	670	383	48	239	1052	96	112	255	351
BLOOMINGTON	58.6	606	48	96	765	1068	96	159	191	255
MADISON	60.0	159	16	64	159	255	*	16	415	*
BICKNELL	58.0	112	*	*	16	255	48	80	48	64
WASHINGTON	58.0	287	16	80	175	590	128	319	48	144
PETERSBURG	58.6	112	16	32	16	207	16	144	144	128
BOONVILLE	58.7	159	16	48	32	255	16	16	16	32

\* -AT THE LEVEL OF AGGREGATION USED IN LAND USE MEASUREMENT THIS LAND USE WAS NOT OBSERVED.



Table 36. Land Use Measurements for 1970 Sample Cities in Acres

City Name	Date of Airphotos	Intra-City Highways	Industrial	Commercial	Institutional	Residential	Other Transportation	Agriculture	Water & Forest	Vacant
MICHIGAN CITY	65.6	1578	670	175	1148	1977	733	1610	7684	1626
LONG BEACH	65.6	64	*	16	95	144	*	*	335	80
ELKHART CITY	65.7	1626	893	191	542	2631	526	957	1531	1068
GOSHEN	65.7	622	430	112	383	1515	159	1706	813	446
KNOX	65.6	271	32	16	112	319	144	590	128	287
FOKLER	71.6	191	*	16	32	223	*	128	*	144
HARTFORD	69.9	319	159	48	144	622	48	128	223	159
PORTLAND	69.9	303	128	16	191	510	96	223	112	207
COVINGTON	72.8	112	*	16	48	303	32	96	80	*
CRAWFORDSVILLE	71.6	446	287	112	191	1068	64	638	590	112
UPLAND	69.9	96	16	48	144	159	48	96	64	96
NEW CASTLE	69.9	781	223	80	303	1275	128	175	144	415
RICHMOND	69.9	1180	574	686	813	2685	271	1020	829	670
CONNERSVILLE	69.5	590	223	96	415	1226	239	399	526	702
COLUMBUS	68.6	2391	702	303	813	2981	2933	11447	3491	1690
BLOOMINGTON	67.9	645	159	128	1323	1610	64	16	207	335
MADISON	68.8	510	175	64	367	1116	415	638	1817	399
BICKNELL	66.8	207	*	*	16	319	16	*	49	48
WASHINGTON	66.8	558	32	80	112	829	112	96	239	351
JASPER	66.7	542	207	144	191	622	16	765	415	335
FORTBRANCH	66.6	80	16	16	16	223	*	*	*	48
BOONVILLE	66.8	255	32	32	32	415	16	96	128	96
BLOOMFIELD	67.8	191	32	*	*	287	16	48	112	48
DELPHI	71.9	159	32	32	112	207	32	32	128	32

\* - AT THE LEVEL OF AGGREGATION USED IN LAND USE MEASUREMENTS THIS LAND USE WAS NOT MEASURED.





publications (3, 4, 5, 6) and were listed in Tables 37 to 40. Median family income for 1940 was missing. This information was not published and a written inquiry to the Census Bureau showed that the 1940 income data was not collected for cities under 100,000 population (50). Also research in Indiana publications within the state of Indiana failed to reveal this information.

Column 3 of these tables shows the total employment within the city. The variable, employment per 1,000 persons (X5), was used in the path model as proxy for total employment. The reason for normalization of employment by population was to make employment data for different cities comparable.

### Statistical Analysis and Results

The statistical studies were analyzed in four steps: First the path model of Figure 8 was calibrated separately with 1940, 1950, 1960 and 1970 data. Second the validity of the calibrated models were tested for their ability to predict the dependent variables. Third the four calibrated models were qualitatively compared to see if the four models could be combined in a single overall model. In the last step the final calibrated models were used to determine total (direct and indirect) effect of independent variables on the dependent variables.

### Calibration of the Path Models

The path model to be calibrated was represented by the nine recursive regression equations 6a to 6i, discussed on pages 30 to 38. Each of these nine equations were represented by the following general equation:



Table 37. Census Data for 1940 Sample Cities

City Name	Population	Total Employment	Family Median Income
VALPARAISO	8736	3202	N.A.
KENDALLVILLE	5431	2006	N.A.
WARSAW	6378	2346	N.A.
HUNTINGTON	13903	4609	N.A.
KOKOMO	33795	11345	N.A.
HARTFORD CITY	6946	2393	N.A.
LAFAYETTE	28798	10467	N.A.
ATTICA	3760	1257	N.A.
FRANKFORT	13706	4600	N.A.
MUNCIE	49720	18138	N.A.
RICHMOND	35147	13216	N.A.
GREENCASTLE	4872	1765	N.A.
CONNERSVILLE	12898	4746	N.A.
SHELBYVILLE	10791	3848	N.A.
GREENSBURG	6065	2026	N.A.
BATESVILLE	3065	1085	N.A.
BLOOMINGTON	20870	6080	N.A.
LINTON	6263	1539	N.A.
BEDFORD	12514	3249	N.A.
VINCENNES	18228	6066	N.A.
JASPER	5041	1824	N.A.

N.A. - NOT AVAILABLE



Table 38. Census Data for 1950 Sample Cities

City Name	Population	Total Employment	Family Median Income
ROCHESTER	4673	1888	2350
MONTICELLO	3467	1468	2713
LOGANSFORT	21031	8662	2720
KOKOMO	38672	15905	3060
HARTFORD	7253	2883	2674
PORTLAND	7064	2863	2875
FRANKFORT	15028	5813	2909
TIPTON	5633	2102	2195
ALEXANDRIA	5147	2023	2862
ANDERSON	46820	21234	2802
NEW CASTLE	18271	7464	3153
RICHMOND	39539	16865	2591
GREENCASTLE	6888	2359	688
SHELBYVILLE	11734	5206	2639
FRANKLIN	7316	2784	2429
BLOOMINGTON	28163	10017	1103
AURORA	4780	1804	2697
BEDFORD	12562	4956	2759
BICKNELL	4572	1474	2521
VINCENNES	18831	7215	2513
PETERSBURG	3035	1025	2042
JASPER	5215	2080	2897
TELL CITY	5735	2343	2718
NAPANEE	3393	1318	2659



Table 39. Census Data for 1960 Sample Cities

City Name	Population	Total Employment	Family Median Income
MICHIGAN CITY	36653	12990	6163
GOSHEN	13718	5864	6016
PLYMOUTH	7558	2958	5781
LOGANSPOUT	21106	8098	5572
BERNE CITY	2644	1112	6113
LAFAYETTE	12680	4974	8068
TIPTON	5604	2252	5771
ELWOOD	11793	4270	5638
COVINGTON	2759	993	5648
ANDERSON	49061	20160	5995
CHESTERFIELD	2588	1039	7441
LEBANON, UEN	9523	3648	5535
NOBLESVILLE	7644	3018	6032
RICHMOND	44149	17343	5583
SHELBYVILLE	14317	5705	5585
MARTINSVILLE	7525	2809	5541
EDINBURG	3664	1235	4746
COLUMBUS	20778	8201	6094
BLOOMINGTON	31357	12182	5448
MADISON	10448	3805	4854
BICKNELL	3878	1263	4702
WASHINGTON	10846	3860	4808
PETERSBURG	2939	956	4431
BOONVILLE	4801	1742	4666
LAWRENCEBURG	5004	1811	5638
SCOTTSBURG	3810	1453	4928





Table 40. Census Data for 1970 Sample Cities

City Name	Population	Total Employment	Family Median Income
MICHIGAN CITY	39369	15413	11240
LONG BEACH	2740	983	20997
ELKHART CITY	43152	18603	12812
GOSHEN	17171	7825	11181
KNOX	3519	1359	8415
FOWLER	2643	1061	10616
HARTFORD	8207	3139	9809
PORTLAND	7115	3112	9537
COVINGTON	2641	1090	10116
CRAWFORDSVILLE	13842	5542	10248
UPLAND	3203	1246	9080
NEW CASTLE	21215	7963	9894
RICHMOND	43999	17765	10489
CONNERSVILLE	17604	6764	10066
COLUMBUS	27141	11078	11948
BLOOMINGTON	42890	16669	10773
MADISON	13081	5035	10512
BICKNELL	3717	1290	7536
WASHINGTON	11358	3042	9122
JASPER	8641	3495	11116
FORTBRANCH	2535	987	9791
BOONVILLE	5736	2108	10332
BLOOMFIELD	2565	1059	10099
DELPHI	2582	1180	10166



$$X_i = \beta_0 + \sum_{j=1}^n \beta_j X_j + e$$

The calibration process used the values of the observed data to estimate the empirical coefficients,  $b_0, b_1, \dots, b_n$  (corresponding to the theoretical coefficients  $\beta_0, \beta_1, \dots, \beta_n$ ).

The regression coefficients were estimated by the ordinary least square regression analysis. The computations were performed on the CDC 6500 computer at Purdue University using the "backward elimination" procedure (52) of the statistical package for the social sciences computer program (SPSS) (53). The program output simultaneously produces regression coefficients and path coefficients. The regression coefficients which upon verification were used for prediction purposes were calculated by use of the raw observed data and the path coefficients were calculated with the standardized data (mean = 0 and variance = 1).

The "backward elimination" procedure first develops a regression equation with all the hypothesized variables in the model then:

1. Calculates partial F values for each of the model variables assuming that the remainder of the variables were already in the model.
2. Compares the lowest of the partial F's with a critical F value computed for a specified significance level (F critical was chosen for  $\alpha = 0.05$ ).
3. As long as the lowest partial F value was less than the critical F value, the corresponding variable was dropped from the regression equation and step 2 was repeated. Otherwise the regression equation was accepted.



### Hypothesis Testing of Regression Coefficients

Each of the variables in the final regression equations were tested for significant contribution to the equation after due consideration to the influence of other variables already in the equation. If an independent variable significantly affects the regression equation, its corresponding regression coefficient was non zero. The null hypothesis ( $H_0$ ) assumes that the regression coefficient was zero and the alternative hypothesis ( $H_1$ ) assumes that the regression coefficient was non-zero. If the null hypothesis was rejected then the alternative hypothesis was not rejected.

Rejection of the null hypothesis, in part, depends on the level of significance used in the hypothesis testing. In this project a significance level of five percent was used ( $\alpha$  = type I error = 0.05); which means there was a five percent chance of erroneously rejecting the null hypothesis when it should have been accepted. The specific test procedure for each regression coefficient follows:

1.  $H_0$ , null hypothesis:  $b_i = 0$   
 $H_1$ , alternative hypothesis:  $b_i \neq 0$
2. Estimated statistic, Partial F:  $F(K, n-K-1) = \frac{(R_K^2 - R_{K-1}^2)/1}{(1-R_K^2)/(n-K-1)}$

where:

$n$  = number of observations

$K$  = number of independent variables

$R^2$  = coefficient of multiple determination

3. Critical F Statistic:  $F(K, n-K-1, 1-\alpha)$  where  $K$  and  $n-K-1$  are appropriate degrees of freedom.



4. Reject the null hypothesis if:

estimated statistic > critical statistic

#### Results of 1940, 1950, 1960 and 1970 Model Calibrations

The results of model calibrations for the 1940, 1950, 1960 and 1970 data periods are shown in Tables 41 to 44 and Figures 18 to 21. The tables present the calibrated path model with unnormalized variables and regression coefficients. The equation in these tables after verification were used for predictive purposes. The figures show the same calibrated model, but with the normalized variables and path coefficients. These path coefficients were used to obtain the direct and indirect effect of an independent variable on an appropriate dependent variable.

The results of the calibration process show that generally the dependent variables were adequately explained by the models, but there were differences between the level of explanation of the four models. In the discussion of the ability of the equations to explain variations of the dependent variables it was noted, that there are two distinct sets of variables in the transportation-city development path model (endogenous and exogenous). The purpose of the hypothesized general model was to explain variations of residential, institutional, vacant, commercial, and industrial land use (endogenous variables), while the purpose of inclusion of city locations, inter-city accessibility, intra-city accessibility, employment, median family income, and population (exogenous variables) in the model was to assist in explanation of variation of land uses.





It was expected the equations having the endogenous variables as dependent would show a higher level of explanation of dependent variables (have higher  $R^2$ ) than the equations having the exogenous variables as dependent. The latter equations were required for calculations of total effects coefficients, but a high level of explanation was not expected from them.

#### 1940 Path Model

The original hypothesized model of Figure 6 has 54 links, but because of unavailability of median family income for 1940 the number of links in the 1940 hypothesized model was reduced to 44. Figure 18 shows the calibrated 1940 model to have retained 12 links at a significance level of 95 percent. These 12 links explain most of the variations of the dependent variables as shown by relatively high coefficients of determinations ( $R^2$ ) associated with their respective regression equation shown in Table 41. The calibrated model does not describe the variations of each of the dependent variables equally. The following discussion pertains to the model's ability to explain each of the variables and reasonableness of the associated regression equations.

#### Residential Land Use

About ninety-three percent of the variations in the acres of land allocated for residential land use in 1940 sample cities was explained by the calibrated models in agreement with the hypothesized assumption. Residential land use was positively correlated with city location and city population. The residual of this equation had no pattern; although, the absolute amount of the error becomes larger for larger cities.



Table 41. 1940 Path Model with Unnormalized Variable and Regression Coefficients

Dependent Variable	Description of Individual Equations (Numbers in parentheses under the regression coefficients are the level of significance, $\alpha$ , for the corresponding regression coefficient)	Statistical Measures for Each Equation		
		Coefficients of Multiple Determination ( $R^2$ )	Mean Square Error ( $S^2$ )	Level of Significance ( $\alpha$ )
Residential Land Use	83.05 + 1.028 City Location + 0.0364 Population (0.000)	93.44	16,629	0.00
Institutional Land Use	29.24 - 1.635 Commercial Land Use + 0.0109 Population (0.000)	77.47	4,831	0.00
Vacant Land Use	91.49 - 1.920 Commercial Land Use + 0.552 Intra-City Accessibility (0.000)	74.87	81,660	0.00
Commercial Land Use	24.77 - 0.375 Industrial Land Use + 0.00341 Population (0.004)	49.10	766	0.00
City Population	-374.3 + 34.308 Industrial Land Use + 31.855 Intra-City Accessibility (0.028)	91.11	15.43x10 <sup>6</sup>	0.00
Median Family Income	Income Data Not Available for 1940			
Employment per 1000 Persons	246.57 + 0.002326 Inter-City Accessibility (0.030)	22.36	10,171	0.03
Industrial Land Use	-5.44 + 0.214 Intra-City Accessibility (0.000)	52.31	3,947	0.00
Intra-City Accessibility	No Significant Equation			



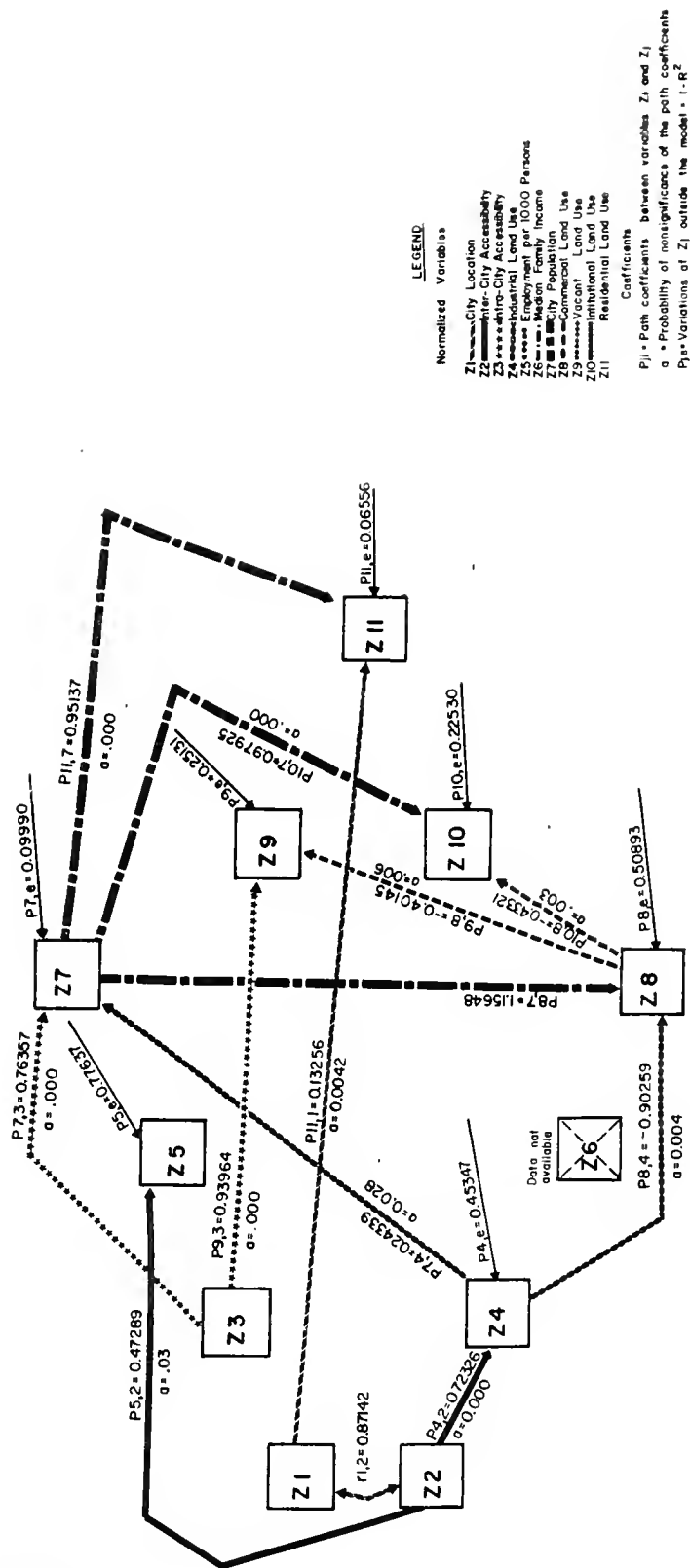


Figure 18. Path Coefficients of Calibrated 1940 Transportation City Development Model.



### Institutional Land Use

About seventy-seven percent variations of institutional land use was explained by the 1940 model. Institutional land use was positively correlated with city population and negatively correlated with commercial land use. Positive correlation of institutional land use with city population is in agreement with the hypothesized path model, but its negative correlation with commercial land use was in disagreement with the hypothesized model and not readily explained.

### Vacant Land Use

About seventy-five percent of variation in vacant land use in the sample cities was explained by the model. Vacant land use was positively related to acres of land used for city streets and negatively related to commercial land use. The dependent variable (vacant land use) as used in this study was in reality the land reserved for future use. It would seem reasonable that highway and street construction would increase the amount of land reserved for future use. The negative influence of commercial land use on the dependent variable was reasonable because more commercial land use signifies a higher degree of land use and less availability of vacant land within the confines of the study where the size of each city was limited to the land within the city limits.

### Commercial Land Use

Only about forty-nine percent of variation in the commercial land use was explained by the two independent variables in the equation. The industrial land use was negatively related with commercial land use. This negative relationship was contrary to prior belief and not readily explainable. Furthermore, the results obtained from the rest





of the data (see Table 42) shows a positive relation between the industrial land use and commercial land use. This leads to the belief that either the 1940 sample data does not represent the inference space adequately or the possibility of significant differences in industrial-commercial land use patterns before and after World War II. The second variable, population, was positively related to the commercial land use, which was in agreement with hypothesized assumptions.

#### City Population

City population was not truly a dependent variable, but was entered in the model to help explain variations in the land uses. Ninety-one percent of the variation in the population was explained by the industrial land use and the city's street system. Positive correlation of industrial land use and city streets with population is in agreement with the assumptions of the hypothesized path model.

#### Employment Per 1000 Persons

The model explains only twenty-two percent of variation in city employment by variation of inter-city accessibility. Such a low  $R^2$  by itself would not necessarily be unsatisfactory because employment was not really a dependent variable, but it was in the model to help to explain variations in land uses. Unfortunately in this model there were no significant relationships between employment and any of the land uses; as a result the employment variable as used in this model does not serve any useful purpose.



### Industrial Land Use

Fifty-two percent of variations in the industrial land use was explained by the variations in the street system. There was a linear trend on the residuals of this regression equation. The reason for the trend was that one or more significant independent variables were missing from this equation. Nevertheless it was significant that fifty percent of variations in industrial land use was explained by the land used for the city streets.

### Intra-City Accessibility

There was no satisfactory regression equation of intra-city accessibility with city location and inter-city accessibility. Therefore, the original hypothesis statement that location and inter-city accessibility influence intra-city accessibility was not verified.

### Location and Inter-City Accessibility

These two variables were highly correlated (correlation coefficient,  $r_{1,2} = 0.87$ ) the explanation of this correlation was outside the scope of the model.

### 1950 Path Model

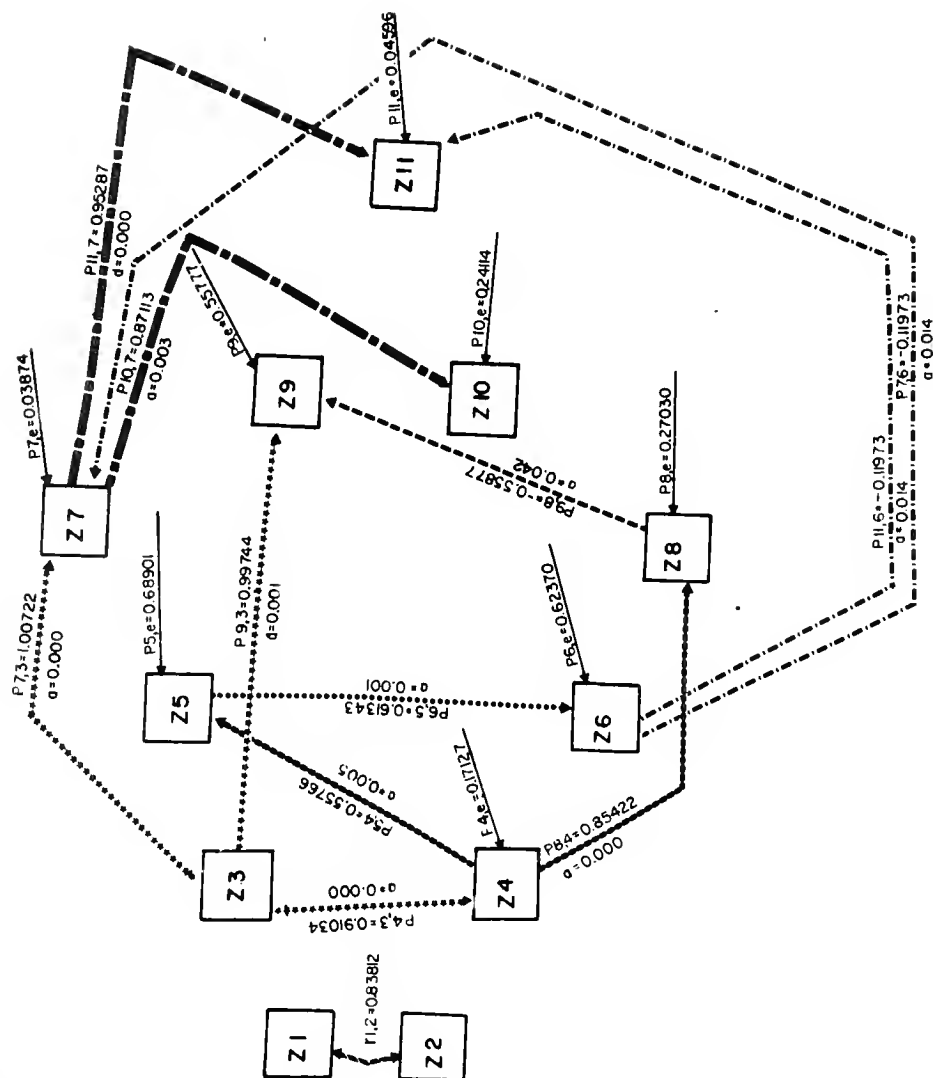
From the original 54 causal links in the generalized transportation-city development path model the relationship represented by 11 of these links was confirmed. Table 42 and Figure 19 show these significant links. The 1950 model explains variations of residential, institutional, commercial, and industrial land use adequately, but the variations of the vacant land use was not explained adequately.



Table 42. 1950 Path Model with Unnormalized Variables and Regression Coefficients

Dependent Variable	Description of Individual Equations (Numbers in parenthesis under the regression coefficients are the level of significance, $\alpha$ , for the corresponding regression coefficient)	Statistical Measures for Each Equation			
		Coefficients of Multiple Determination ( $R^2$ )	Mean Square Error ( $S^2$ )	Level of Significance ( $\alpha$ )	
Residential Land Use	$-132.11 + 0.0982 \text{ Median Family Income} + 0.0384 \text{ City Population}$ (0.03) (0.00)	95.40	13,234	0.00	
Institutional Land Use	$23.59 + 0.00797 \text{ Population}$ (0.00)	75.89	3,417	0.00	
Vacant Land Use	$33.34 - 0.698 \text{ Commercial Land Use} + 0.339 \text{ Intra-City Accessibility}$ (0.04) (0.00)	44.22	3,977	0.00	
Commercial Land Use	$16.80 + 0.378 \text{ Industrial Land Use}$ (0.00)	72.97		0.00	
City Population	$4431.05 - 2.656 \text{ Median Family Income} + 54.013 \text{ Intra-City Accessibility}$ (0.01) (0.00)	96.12	$6.863 \times 10^7$	0.00	
Median Family Income	$-1822.5 + 11.109 \text{ Employment Per 1000 Persons}$ (0.00)	37.63	214,314	0.00	
Employment per 1000 Persons	$380.96 + 0.121 \text{ Industrial Land Use}$ (0.01)	31.10	722	0.01	
Industrial Land Use	$-64.98 + 0.561 \text{ Intra-City Accessibility}$ (0.00)	82.87	3,820	0.00	
Intra-City Accessibility	No Significant Equation				









### Residential Land Use

Ninety-five percent of variation of residential land use in the 1950 sample cities was explained by variation of city population and median family income. Correlation of higher income and higher residential land use was reasonable because generally the more affluent population live in residential units with the larger lots.

### Institutional Land Use

About seventy-six percent of variation in the 1950 institutional land use was explained by the city population. As institutional land use was there to serve the people this equation was intuitively correct and in agreement with prior assumptions.

### Vacant Land Use

Only forty-four percent of variations in the vacant land was explained by the 1950 model. Vacant land use varies negatively with commercial land use, and positively with the city's street system.

### Commercial Land Use

Seventy-three percent of variations in the commercial land use was explained by the industrial land use. This was in agreement with the assumptions that commercial development follows the industrial development to serve the industry and the workers.

### City Population

Ninety-six percent of the variation in the city population was explained by the 1950 model. City population was negatively related to median family income and positively related to the city's street



system. The negative relation between income and population implies that, other conditions being equal, city's with a higher median family income have a smaller number of inhabitants.

#### Median Family Income

Only thirty-eight percent of the variation in the median family income was explained by the variation among employment per 1000 persons in the 1950 model. But this low  $R^2$  was not unsatisfactory, because the purpose of the model was not to predict median family income, but to deploy this variable to explain variations in land uses. The positive correlation of median family income with employment per 1000 persons was in agreement with the hypothesized assumptions.

#### Employment Per 1000 Persons

Only thirty-one percent of the variation in the employment rate was explained by the variation in the industrial land use in the 1950 model. The correlation of employment rate and industrial land use was intuitively correct and in agreement with prior assumptions. Evidently the employment rate depends on other factors not in the equation, such as the number of workers who commute to other cities for work. Nevertheless this relationship was useful because its purpose was to assist in the explanation of the variations of the land use and not to predict employment rate.

#### Industrial Land Use

About eighty-three percent of the variation in the industrial land use was explained by the variation in the city's street system by the 1950 model.



### Intra-City Accessibility

There were no satisfactory regression equations to explain the variations of intra-city accessibility with city location and inter-city accessibility.

### Location and Inter-City Accessibility

These two variables were highly correlated (correlation coefficient,  $r_{1,2} = 0.84$ ). The explanation of this correlation was outside the scope of the hypothesized path model.

### 1960 Path Model

The original hypothesized model of Figure 6 has 54 links. The calibrated 1960 model as presented in Table 43 and Figure 20 confirms 19 of these links. These 19 links adequately explain variations of residential, institutional, vacant, and industrial land use, but do not adequately explain the variations in the commercial land use.

### Residential Land Use

Ninety-eight percent of the variation in the residential land use was explained by the 1960 model. Residential land use was negatively related with employment per 1000 persons and institutional land use; and positively related with city location, intra-city accessibility, and population. The negative correlation of residential land use with employment and institutional land use was contrary to original assumptions, and their implication was counter-intuitive. Negative correlation of employment per 1000 persons and institutional land use with residential land use implies that the city's with higher per



Table 43. 1960 Path Model with Unnormalized Variables and Regression Coefficients

Dependent Variable	Description of Individual Equations (Numbers in parenthesis under the regression coefficients are the level of significance, $\alpha$ , for the corresponding regression coefficient)	Statistical Measures for Each Equation			
		Coefficients of Multiple Determination ( $R^2$ )	Mean Square Error ( $S^2$ )	Level of Significance ( $\alpha$ )	
Residential Land Use	721.79 + 0.923 City Location - 1.74 Employment per 1000 (0.01) (0.048)				
	- 0.486 Institutional Land Use + 0.474 Intra-City Accessibility (0.004) (0.008)				
	+ 0.0377 Population (0.000)	98.26	6,583	0.00	
Institutional Land Use	-41.95 - 0.923 Commercial Land Use - 1.125 Industrial Land Use (0.00) (0.00)				
	+ 0.567 Vacant Land Use + 0.0241 Population (0.00) (0.00)	96.50	55,533	0.00	
Vacant Land Use	-70.24 - 0.0192 Population + 1.410 Intra-City Accessibility (0.00) (0.00)	86.93	10,413	0.00	
Commercial Land Use	3.45 + 0.144 Intra-City Accessibility (0.00)	41.79	3,367	0.00	
City Population	-16947.0 + 79.786 Employment per 1000 Persons (0.04)				
	+ 38.727 Intra-City Accessibility - 0.211 Inter-City Accessibility (0.00) (0.03)	90.80	17,849,584	0.00	
Median Family Income	1115.5 + 4.790 City Location + 11.949 Employment per 1000 (0.04) (0.04)	48.48	381.124	0.00	
Employment per 1000	299.18 + 0.00116 Inter-City Accessibility (0.01)	27.08	538	0.00	
Industrial Land Use	-57.14 + 0.444 Intra-City Accessibility (0.00)	82.26	4,982	0.00	
Intra-City Accessibility	-522.06 + 0.0131 Inter-City Accessibility	21.87	91,491	0.02	





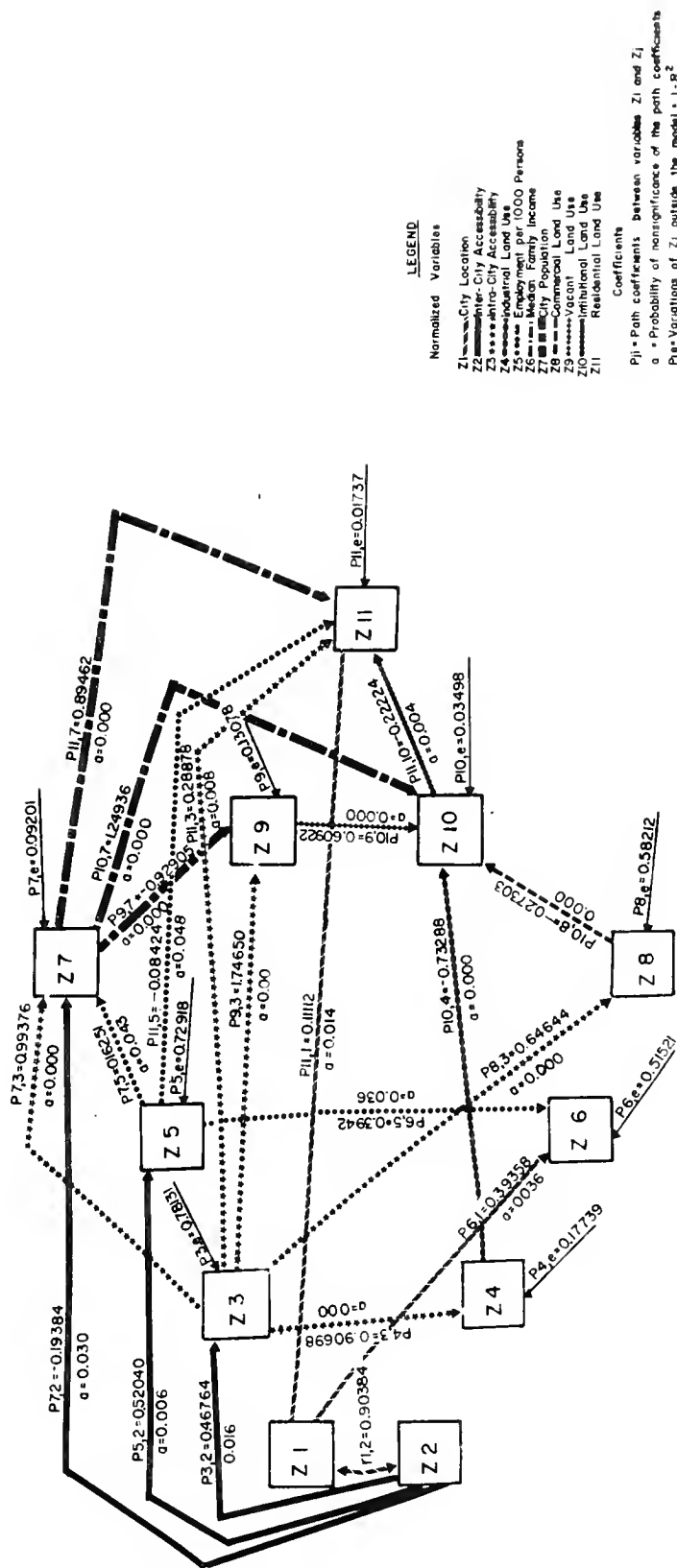


Figure 20. Path Coefficients of Calibrated 1960 Transportation City Development Model.



capita employment and higher institutional land use allocate smaller acreage for residential land use.

#### Institutional Land Use

About ninety-six percent of the variation in the institutional land use was explained in the 1960 model. Institutional land use was positively related with population and vacant land; and negatively related with commercial and industrial land use. The implication of negative correlation was that cities with more industrial and commercial land use allocate less land for institutional purposes.

#### Vacant Land Use

About eighty-seven percent of the variation of vacant land use was explained by the 1960 model. Vacant land use was positively related with intra-city accessibility and negatively with population. This equation seems reasonable because a more extensive street system creates more vacant land (land reserved for future use) and a larger population creates pressures for deployment of the reserved land.

#### Commercial Land Use

Only about forty-two percent of the variation in the commercial land use was explained by the 1960 model. Commercial land use was positively related with the intra-city accessibility. This positive correlation of commercial land use and intra-city accessibility seems reasonable.



### City Population

About ninety percent of the variation in the city population was explained by the 1960 model. City population was positively correlated with employment and intra-city accessibility, and negatively correlated with inter-city accessibility. The negative correlation of city population and inter-city accessibility would seem to be counter-intuitive. Inter-city accessibility was a function of the region wide highway system and the population of all other cities in the region. The negative correlation between the population and inter-city accessibility may imply that when the smaller city was located near a larger city; the larger city was negatively influencing the growth of the smaller city.

### Median Family Income

Only forty-eight percent of the variation in the median family income was explained by the 1960 model. Median family income was positively related with the city location and employment per 1000 persons. The positive correlation of median family income and employment rate intuitively seems reasonable. The positive correlation of income and city location implies that for a given employment rate the cities in the northern part of Indiana have a higher median family income than cities in southern Indiana.

### Employment Per 1000 Persons

Only twenty-seven percent of the variation in the employment rate was explained in the 1960 model. Employment was positively correlated with inter-city accessibility which seems reasonable.



### Industrial Land Use

About eighty-two percent of the variation in the industrial land use was explained by the 1960 model. Industrial land use was positively related with intra-city accessibility, which was in agreement with prior assumptions.

### Intra-City Accessibility

Only twenty-two percent of the variation in the intra-city accessibility was explained by the 1960 model. Intra-city accessibility was positively related with inter-city accessibility, which implies that, other conditions being equal, cities with better access to other cities also have a better street system.

### Location and Inter-City Accessibility

These two variables were highly correlated (correlation coefficients,  $r_{1,2} = 0.90$ ). The explanation of this correlation was outside the scope of the hypothesized path model.

### 1970 Path Model

Calibration of the path model of Figure 7 with the 1970 data confirms 18 of the 54 original hypothesized links. These 18 links adequately explain the variations of all 1970 land uses with the exception of commercial land use which was not adequately explained. Table 44 and Figure 21 contain the calibration results of the 1970 path model.





Table 44. 1970 Path Model with Unnormalized Variables and Regression Coefficients

Dependent Variable	Description of Individual Equations (Numbers in parenthesis under the regression coefficients are the level of significance, $\alpha$ , for the corresponding regression coefficient)	Statistical Measures for Each Equation			
		Coefficients of Multiple Determination ( $R^2$ )	Mean Square Error ( $S^2$ )	Level of Significance ( $\alpha$ )	
Residential Land Use	-1516.8 - 4.150 City Location + 1.368 Commercial Land Use (0.01) (0.00)				
	+ 0.0164 Inter-City Accessibility + 0.738 Intra-City Accessibility (0.00) (0.00)				
	+ 0.0174 Population (0.00)				
Institutional Land Use	-927.90 + 2.26 Employment per 1000 + 0.526 Vacant Land Use (0.00) (0.00)	97.70	22,920	0.00	
	+ 0.0299 Population - 1.363 Industrial Land Use (0.00) (0.00)				
Vacant Land Use	-47.40 + 0.756 Intra-City Accessibility (0.00)	94.13	9,455	0.00	
Commercial Land Use	-5.03 + 0.00734 Population (0.00)	89.29	24,349	0.00	
City Population	4771 + 46.228 Industrial Land Use (0.00)	54.45	10,161	0.00	
Median Family Income	No Significant Equation	68.59	6.806x10 <sup>7</sup>	0.00	
Employment Per 1000 Persons	391.37 - 0.0616 Intra-City Accessibility + 0.176 Industrial Land Use (0.05) (0.02)	26.61	1,219	0.04	
Industrial Land Use	-327.05 + 0.368 Intra-City Accessibility + 0.00307 Inter-City (0.00) (0.00)				
	Accessibility	88.52	8,366	0.00	
Intra-City Accessibility	-3206.28 - 8.922 City Location + 0.0370 Inter-City Accessibility (0.04) (0.02)	23.96	282,550	0.06	



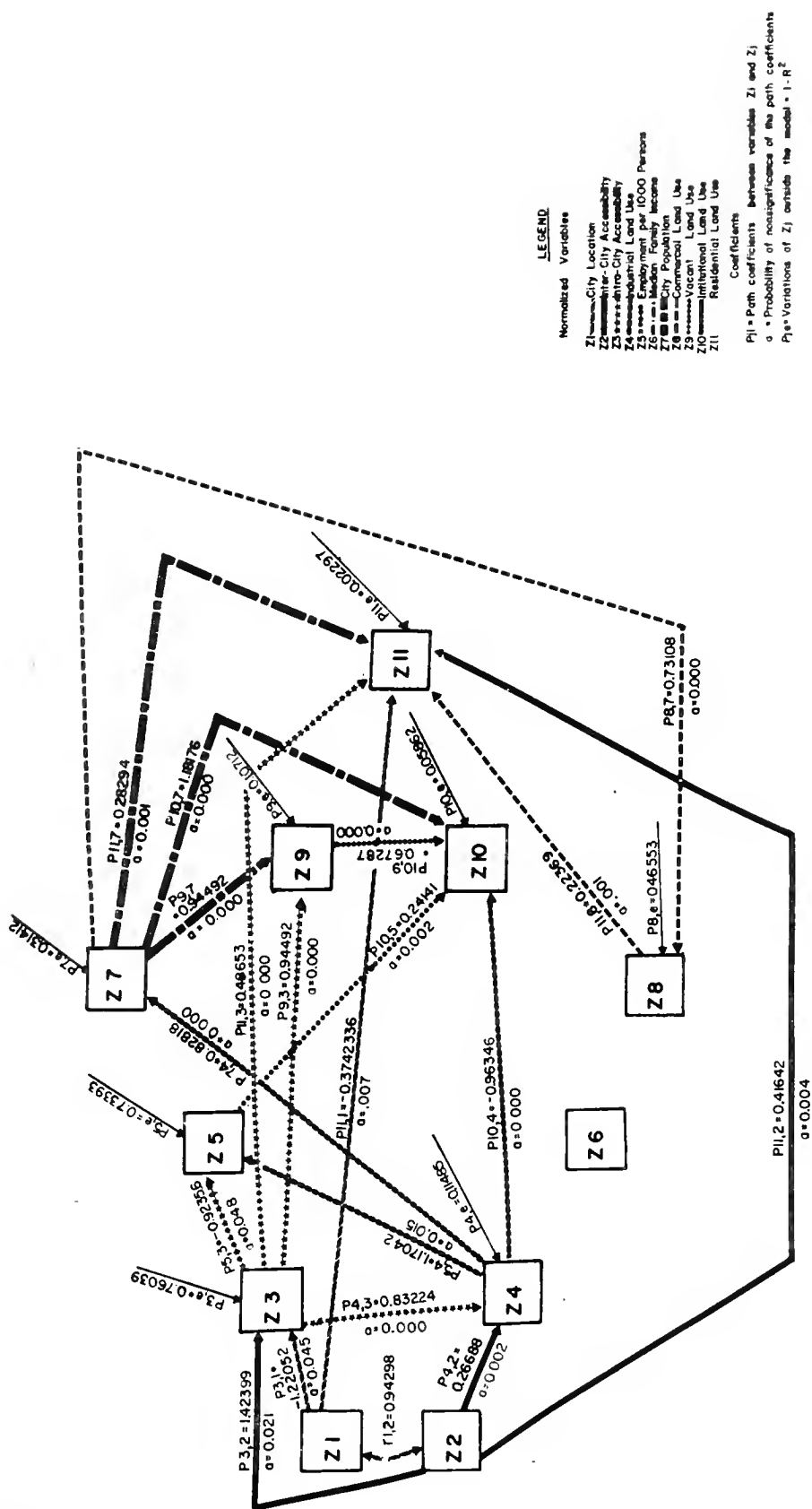


Figure 21 Path Coefficients of Calibrated 1970 Transportation City Development Model.



### Residential Land Use

About ninety-eight percent of the variations of residential land use was explained by the 1970 model. Residential land use was positively correlated with commercial land use, inter-city accessibility, intra-city accessibility, and city population; and negatively correlated with city location. The positive correlations were in agreement with the hypothesized model and intuitively reasonable. The negative correlation with the city location implies that, other conditions being equal, the southern Indiana cities allocate less land for residential land use than cities in northern Indiana.

### Institutional Land Use

About ninety-four percent of the variation in the institutional land use was explained by the 1970 model. Institutional land use, in agreement with the hypothesized model, was positively correlated with employment per 1000 persons, vacant land use and population, but contrary to prior assumption was negatively correlated with industrial land use. This negative correlation may imply that, other conditions being equal, highly industrialized cities use less of their land for public service purposes.

### Vacant Land Use

About eight-nine percent of the variation of the vacant land use was explained by the 1970 calibrated model. Vacant land use was correlated positively with intra-city accessibility, which was in agreement with the hypothesized assumptions.



### Commercial Land Use

Only fifty-three percent of variations in the commercial land use was explained by the 1970 calibrated model. But positive correlation of commercial land use and population was in agreement with the hypothesized assumptions.

### City Population

About sixty-eight percent variations of city population was explained by the calibrated 1970 model. The positive correlation of city population with industrial land use was in agreement with the hypothesized assumptions.

### Median Family Income

At 95 percent significance level ( $\alpha = 0.05$ ) there were no significant regression equation developed to explain the variations of median family income. Therefore no equation for median family income was shown in the Table 44. But, the following equation:

$$\text{Median Family Income} = 10,526.9 + 12.08 \text{ City Location}$$

which explains fifteen percent of variations in median family income was significant at the 94.2 percent ( $\alpha = 0.058$ ). This equation implies, all other conditions being equal, a higher median family income in the northern Indiana cities than southern Indiana cities.

### Employment Per 1000 Persons

Only about twenty-seven percent of the variation in the employment rate was explained by the 1970 calibrated model. Employment rate





was correlated negatively with intra-city accessibility and positively with industrial land use. The negative correlation of employment rate with intra-city accessibility was contrary to the hypothesized assumption and in disagreement with the results obtained from the 1940, 1950, and 1960 calibrated models. The positive correlation, of employment rate and industrial land use was in agreement with the hypothesized assumptions.

#### Industrial Land Use

Eighty-eight percent variation of industrial land use was explained by the 1970 calibrated model. Industrial land use was positively correlated with both intra-city accessibility and inter-city accessibility. These correlations were in agreement with hypothesized assumptions and may mean that better inter and intra-city accessibility positively influence industrial development.

#### Intra-City Accessibility

Only twenty-four percent variation of the intra-city accessibility was explained by the 1970 model. The intra-city accessibility was positively correlated with inter-city accessibility, which may mean that cities with higher inter-city accessibility also have a higher intra-city accessibility. The intra-city accessibility was negatively related with city location, which implies northern Indiana cities deploy less land for their street system compared with cities in southern Indiana.



### Location and Inter-City Accessibility

These two variables were highly correlated (correlation coefficient,  $r_{1,2} = 0.94$ ). The explanation of this correlation was outside the scope of the hypothesized path model.

### Ability of Calibrated Path Models to Predict Land Use Variables

General path model of Figure 6 in the previous section was calibrated separately with the collected data for 1940, 1950, 1960 and 1970. An analysis was made to determine whether the calibrated models were capable of predicting land uses within the 95 percent confidence interval or not. For this purpose it was necessary to have available data for cities within the inference space which were not used in the calibration of the model under investigation. But in the calibration process of the last section it was desirable to use all the available information to obtain the best possible models, and there were no unused data to test the calibrated models of Tables 41 to 44.

To overcome this difficulty, three cities were removed from the data set for each of the four decades and the remainder of the data were used to recalibrate the parameters of the regression equations. The three cities removed from each data set were chosen randomly; one each from the three different city size groups. The new calibration process for each regression equation did not begin with the list of independent variables hypothesized by the original model of Figure 6, but used the significant independent variables determined by the previous calibration process to estimate the new parameters associated with the reduced data set.



The recalibrated regression equations were then used to estimate the 95 percent confidence interval for the land uses of the three cities not used in their calibration process. If the measured land uses fell in the 95 percent confidence interval predicted by the recalibrated models then the calibrated models were valid to predict land use for other cities in the inference space at the 95 percent significant level.

Estimation of the standard error for the predicted mean value of a dependent variable depends on the specified values of the set of independent variables and was given by the following equation (54) in the matrix notation:

$$S_{Y(\text{mean})} = \text{MSE } (X_0' (X'X)^{-1} X_0)$$

where

MSE = mean square error

$X_0$  = set of specified independent variables for which  
standard error is to be estimated

$X$  = set of independent variables used to calibrate the  
regression equation

$X_0', X'$  = transpose of the respective matrix

$(X'X)^{-1}$  = inverse of the  $X'X$  matrix.

The present task requires prediction of each dependent variable based on a single observation for a specified set of independent variables. The predicted value of this dependent variable was given by the  $\hat{Y}$ , but the standard error was larger than given by the above formula, and was given by



$$S_Y^2 = \text{MSE} (1 + X_0'(X'X)^{-1}X_0)$$

As a result, the 95 percent confidence interval for the predicted value of a single observation was given by

$$Y \pm t(0.975; n-P-1)(S_Y)$$

where

$n$  = number of observations used to calibrate the regression equation

$P$  = number of independent variables in the regression equation.

Program CPIYL, prepared by Purdue University statistics department (55), was used to recalibrate all the regression equations, with one of the land uses as the dependent variable and it was used to calculate the standard error, and the 95 percent confidence interval for the three cities not used in the recalibration of the models.

Tables 45 to 48 contain the results from the above analysis for 1940, 1950, 1960, and 1970 data sets. In each table the measured value, the predicted value, the standard error, and the 95 percent confidence interval for each of the five land uses were given for the three cities not used in the recalibration of the regression equations. The results presented in Tables 45 to 48 were conservative in the sense that the original calibrated equations as presented in Tables 41 to 44 (which were used for any other prediction of land uses) were better equations (they have larger degrees of freedom and smaller mean square error) than the recalibrated equations used to produce the models contained in Tables 45 to 48.





Table 45. Testing the 1940 Path Model for Predicting Land Use Variables

Dependent Variable	Cities Not Used in Calibration of Regression Equation											
	Attica				Vincennes				Bloomington			
	Measured	Predicted	Standard Error $S_Y$	95 Percent Confidence Interval	Measured	Predicted	Standard Error $S_Y$	95 Percent Confidence Interval	Measured	Predicted	Standard Error $S_Y$	95 Percent Confidence Interval
Residential Land Use (n=18, P=2)	207*	264	123	0 to 528	414*	730	128	455 to 1004	733*	847	125	581 to 1114
Institutional Land Use (n=18, P=2)	16*	54	88	-132 to 240	112*	108	87	54 to 163	239*	178	86	-4 to 361
Vacant & Unused Land (n=18, P=2)	48*	149	103	-69 to 366	128*	144	101	-72 to 361	383*	287	100	74 to 500
Commercial Land Use (n=18, P=2)	16*	38	35	-35 to 112	80*	54	34	-18 to 127	48*	78	36	2 to 154
Industrial Land Use (n=18, P=1)	16*	17	64	-119 to 154	96*	74	63	-59 to 208	48*	108	63	-26 to 243

Notes: \* The measured land use values fall within the 95 percent confidence interval for one observation.

n = number of observations.

P = number of independent variables in the regression equations.



Table 46. Testing the 1950 Path Model for Predicting Land Use Variables

Dependent Variable	Cities Not Used in Calibration of Regression Equation											
	Portland				Shelbyville				Kokomo			
	Measured Y	Predicted Y	Standard Error S <sub>y</sub>	95 Percent Confidence Interval	Measured Y	Predicted Y	Standard Error S <sub>y</sub>	95 Percent Confidence Interval	Measured Y	Predicted Y	Standard Error S <sub>y</sub>	95 Percent Confidence Interval
Residential Land Use (n=21, P=2)	399*	432	111	199 to 665	367*	583	109	353 to 812	1801*	1621	111	199 to 665
Institutional Land Use (n=21, P=1)	47*	83	56	-34 to 202	80*	124	56	7 to 242	223*	362	62	233 to 492
Vacant & Unused Land (n=21, P=2)	16*	74	63	-58 to 207	64*	86	62	-46 to 218	127*	258	72	105 to 410
Commercial Land Use (n=21, P=1)	32*	35	36	-41 to 111	64*	41	36	-35 to 117	96*	135	38	56 to 215
Industrial Land Use (n=21, P=1)	48*	34	67	-106 to 173	64*	90	66	-49 to 228	303*	369	73	217 to 521

Notes: \* the measured values fall within the 95 percent confidence interval for a single observation.

n = number of observations.

P = number of independent variables in the regression equations.



Table 47. Testing the 1960 Path Model for Predicting Land Use Variables

Dependent Variable	Cities Not Used in Calibration of Regression Equation									
	Chesterfield					Shelbyville				
	Measured	Predicted	Standard Error	95 Percent Confidence Interval	Measured	Predicted	Standard Error	95 Percent Confidence Interval	Measured	Predicted
	Y	Y	S <sub>y</sub>		Y	Y	S <sub>y</sub>		Y	Y
Residential Land Use (n=23, P=5)	175*	159	96	-48 to 357	685*	692	95	492 to 891	1721*	1762
Institutional Land Use (n=23, P=4)	32*	33	54	-81 to 147	144*	142	55	27 to 258	1036*	794
Vacant & Unused Land (n=23, P=2)	32*	32	60	-95 to 159	207*	225	60	100 to 352	1339	713
Commercial Land Use (n=23, P=1)	16*	6	58	-115 to 127	48*	78	58	-41 to 197	96	246
Industrial Land Use (n=23, P=1)	0*	-33	74	-186 to 120	207*	144	73	-6 to 295	446*	561
										88
										378 to 745

Notes: \* the measured values fall within the 95 percent confidence interval for one observation.

n = number of observations.

P = number of independent variables in the regression equations.



Table 48. Testing the 1970 Path Model for Predicting Land Use Variables

Dependent Variable	Cities Not Used in Calibration of Regression Equation											
	Fowler				Madison				Elkhart			
	Measured Y	Predicted Y	Standard Error $S_Y$	95 Percent Confidence Interval	Measured Y	Predicted Y	Standard Error $S_Y$	95 Percent Confidence Interval	Measured Y	Predicted Y	Standard Error $S_Y$	95 Percent Confidence Interval
Residential Land Use (n=21, P=4)	223*	255	166	-99 to 610	1116*	897	168	539 to 1254	2630*	2503	192	2092 to 2915
Institutional Land Use (n=21, P=4)	32*	115	96	-87 to 318	367*	310	92	114 to 506	542*	809	127	539 to 1079
Vacant & Unused Land (n=21, P=1)	144*	90	170	-265 to 445	399*	338	168	-14 to 690	1068*	1206	182	826 to 1587
Commercial Land Use (n=21, P=1)	16*	11	108	-214 to 236	64*	97	106	-124 to 319	191*	346	108	101 to 591
Industrial Land Use (n=21, P=1)	0*	80	87	-103 to 264	175*	137	87	-46 to 320	892	660	98	454 to 867

Notes: \* the measured values fall within the 95 percent confidence interval for one observation.

n = number of observations.

P = number of independent variables in the regression equations.





Tables 45 to 48 show that generally the predicted 95 percent confidence intervals contained the values of corresponding measured values of land uses. Comparison of 60 measurements of land uses with their respective predicted 95 percent confidence interval show that all but three of the measurements fall within the predicted 95 percent confidence intervals. The three measurements which were not predicted at the 95 percent confidence level with the recalibrated equations were the 1960 vacant and commercial land use for Michigan City and the 1970 industrial land use for the city of Elkhart.

#### Qualitative Comparison of the Four Calibrated Models

In the previous section, results obtained from the calibrations of 1940, 1950, 1960 and 1970 models was separately reported. The five equations (obtained from the four models) which explain each of the land uses were compared and any difference between the way each of the four models explain the variations of the land uses was reported to study the feasibility of combining the data for the four models into an overall model. In evaluating the difference between two equations which contain different independent variables to explain the variations of the same dependent variable, it was important to know the contribution of each independent variable to the explanation of the variations of the dependent variable. Assume the two equations each contain three independent variables with only one common independent variable. At first glance, the two equations seem quite different but if the common independent variable contributes toward ninety percent of the variation of the dependent variable, the two



different equations are indeed similar. The elasticity of each independent variable measures the contribution of an independent variable toward explanation of the variations of the dependent variable.

Elasticity,  $e_i = \frac{B_i X_i}{Y}$ , was given as part of the SPSS regression output (56), where:

$B_i$  = unnormalized regression coefficient for the independent variable,

$X_i$  = mean of the  $i$ th variable,  $Y$  = mean of the dependent variable.

Table 49 shows the independent variables that explain variations of each of the five land uses by the four models and the elasticity associated with each of the regression coefficients. This table was utilized to examine independent variables used to explain variations of each of the land uses by the four models.

### Qualitative Comparison of the Four Equations

#### Predicting Residential Land Use

All the four models show that residential land use was positively correlated with the city's population. The influence of city population on residential land use was substantial in the 1940, 1950, and 1960 model, but was only moderate in the 1970 model. City location influenced residential land use positively in the 1940 and 1950 models and negatively in the 1970 model, but elasticities of city location variables were respectively, 0.018, 0.002, and 0.048 which means the contribution of the city location variable in explanation of variation of residential land use was practically negligible. For example, city location explains two percent ( $0.018/0.868$ ) of variation of residential



Table 49. Independent Variables Used to Explain Variation of Land Uses in Individual Models

Dependent Variable	1940 Model			1950 Model			1960 Model			1970 Model		
	Independent Variable	Elasticity	Independent Variable	Independent Variable	Elasticity	Independent Variable	Independent Variable	Elasticity	Independent Variable	Independent Variable	Elasticity	Elasticity
Residential Land Use	City Location	0.018				City Location	City Location	0.002	City Location	City Location	-0.048	
	Population	0.850	Population	Population	0.815	Population	Population	0.787	Population	Population	0.257	
		--	Median Family Income	Income	0.392			--			--	
		--			--	Employment/1000	Employment/1000	-1.033			--	
		--			--	Institutional Land Use	Institutional Land Use	-0.166			--	
Institutional Land Use		--			--	Intra-City Accessibility	Intra-City Accessibility	0.280	Intra-City Accessibility	Intra-City Accessibility	0.438	
		--			--			--	Inter-City Accessibility	Inter-City Accessibility	1.764	
		--			--			--	Commercial Land Use	Commercial Land Use	0.161	
	Commercial Land Use	0.654			--			--			--	
	Population	1.397	Population	Population	0.821	Commercial Land Use	Commercial Land Use	-0.244	Population	Population	1.377	
Vacant Land Use		--			--	Industrial Land Use	Industrial Land Use	1.473	Industrial Land Use	Industrial Land Use	-0.907	
		--			--	Vacant Land Use	Vacant Land Use	0.532	Vacant Land Use	Vacant Land Use	0.656	
		--			--			--	Employment/1000	Employment/1000	2.828	
	Commercial Land Use	0.401	Commercial Land Use	Commercial Land Use	-0.400			--			--	
	Intra-City Accessibility	0.981	Intra-City Accessibility	Intra-City Accessibility	1.049	Intra-City Accessibility	Intra-City Accessibility	2.597	Intra-City Accessibility	Intra-City Accessibility	1.121	
Commercial Land Use		--			--	Population	Population	-1.254			--	
	Industrial Land Use	0.637	Industrial Land Use	Industrial Land Use	0.692			--	Population	Population	1.050	
	Population	1.093			--			--			--	
		--			--	Intra-City Accessibility	Intra-City Accessibility	0.940			--	
	Intra-City Accessibility	1.070	Intra-City Accessibility	Intra-City Accessibility	1.652	Intra-City Accessibility	Intra-City Accessibility	1.518	Intra-City Accessibility	Intra-City Accessibility	1.021	
Industrial Land Use		--			--			--	Inter-City Accessibility	Inter-City Accessibility	1.54	
		--			--			--			--	



land use in the 1940 model. The 1950 model was the only model which contained median family income to explain residential land use. Intra-city accessibility was utilized in both the 1960 and the 1970 models.

In the 1960 model, institutional land use and employment per 1000 persons negatively influence residential land use. The above negative correlations were counter-intuitive. The negative influence of institutional land use with elasticity of 0.166 was negligible (7.3 percent), but the negative influence of employment per 1000 persons was substantial (45.6 percent). Only in the 1970 model was inter-city accessibility and commercial land use influencing the residential land use. The influence of inter-city accessibility with elasticity of 1.764 was substantial (66.6 percent), but influence of commercial land use with elasticity of 0.141 was negligible (5.3 percent).

#### Qualitative Comparison of the Four Equations

##### Predicting Institutional Land Use

The 1950 model contains population as the only independent variable to explain variations in the institutional land use. The population variable also contributes substantially to explanation of the variation of institutional land use in the other three models. In the 1940 model and 1960 model, the commercial land use was negatively correlated with institutional land use. Both the 1960 and the 1970 models use the industrial land use and vacant land use to explain part of the variation of institutional land use. The negative correlation of industrial land use with institutional land use was contrary to the hypothesized assumptions and imply that, other conditions being equal,





the industrialized cities were allocating less land for public service purposes. Employment per 1000 persons was used only in the 1970 model. The elasticity of employment per 1000 persons was 2.828 which means the employment rate explains 49 percent of the variation of institutional land use in the 1970 model. The implication of the results obtained from the 1970 model was that the so called "bedroom" cities or those cities with higher employment rate but fewer industries use more of their land for public services.

#### Qualitative Comparison of the Four Equations

##### Predicting Vacant Land Use

Both the 1940 and 1950 models use commercial land use and intra-city accessibility to explain variations of vacant land use. In both models the influence of commercial land use was negative and the influence of intra-city accessibility was positive. The elasticities of the independent variable in the two equations were about the same. The positive correlation of vacant land use and intra-city accessibility was due to the fact that a larger street system creates more land reserved for the future use, and the negative correlation of vacant land with commercial land use implies decreasing quantities of land reserved for future use at later stages of development. Intra-city accessibility was also used to explain, partially, the variations of vacant land use in the 1960 model. The other variable in the 1960 model was population which negatively influences the vacant land use. The negative influence of population on vacant land use seems reasonable, because to satisfy the needs of population vacant land was converted to the higher



land uses. Intra-city accessibility alone explains all the variations in the 1970 vacant land use model.

#### Qualitative Comparison of the Four Equations

##### Predicting Commercial Land Use

Industrial land use was positively correlated with commercial land use in the 1940 and 1950 models. Population was positively correlated with commercial land use in the 1940 and 1970 models. Intra-city accessibility was contained only in the 1960 model and, by itself, explains all the variations of commercial land use. The models, although different from each other, were intuitively correct and in agreement with the hypothesized assumptions.

#### Qualitative Comparison of the Four Equations

##### Predicting Industrial Land Use

Intra-city accessibility, alone, explains all the variations of industrial land use in the 1940, 1950 and 1960 models. The 1970 model, in addition to the intra-city accessibility, contains the inter-city accessibility to explain variations of the industrial land use.

#### Results of Qualitative Comparisons of the Four Models

Comparison of the variables used to explain variation of land uses by the four models showed that there are significant differences in the way the four models explain the variations of land uses for the four decades and it does not seem feasible to combine the data for the four decades to calibrate an overall model.



### Total Effect of an Independent Variable on a Dependent Variable

The individual path coefficients of Figure 18 to 21 represent change of the dependent variable, measured in standard deviation, as direct result of one standard deviation change in the independent variable. But the independent variables also affect the dependent variable indirectly, through the intermediary independent variables. Direct effect of an independent variable on the dependent variable is biased because often the consideration of indirect effects of the independent variable on the dependent variable may result in a total effect different from the direct effect. In this section the total effect of each of the independent variables for each of the dependent variables, for the 1940, 1950, 1960 and 1970 data sets, was calculated.

#### Procedure for Calculation of Total Effects of Independent Variables on the Dependent Variables for the General Path Model

The principles of calculations of total (direct and indirect) effect of an independent variable on a dependent variable was discussed in pages 29 to 30 for the four variable example path model of Figure 4. The values for total effect of an independent variable on the dependent variable is obtained by writing all possible paths between the two variables, multiplying the path coefficients associated with each link of the path, and finally summing the partial effect coefficient for each of the paths to lead to the total effects of the independent variables on the dependent variables.



The equations used in calculations of total effect coefficients for the hypothesized eleven variable path model are quite lengthy. The path model variable  $Z_i$  affects variable  $Z_j$  through  $Z^{(j-i-1)}$  different paths. For the hypothesized eleven variable model, the equation measuring total effects of variable  $Z_1$  on variable  $Z_{11}$  contains  $Z^{(11-1-1)} = 512$  different terms because variable  $Z_1$  effects variable  $Z_{11}$  thru 512 different paths. Table 50 shows number of possible paths through which any of the independent variables could influence any of the dependent variables. Comparison of generalized path model of Figure 8 with the calibrated models of Figure 18 to 21 shows many of the paths were statistically insignificant. The number of terms in the total effect equations are smaller than the number given in Table 50. Nevertheless it was found easier to solve the general problem for an eleven variable model and to substitute zero for values of insignificant path coefficients, than developing separate solutions for the four models.

The general solution was obtained by first developing the set of all possible paths from each independent variable to each dependent variable. The above set was formed based on the principle that all the possible paths were formed by all possible permutations of 11 numbers with the important constraint that the numbers had to be arranged in ascending order. The set which is presented in the Appendix B was used in lieu of the equations of total effect because once all the possible paths between the independent variables and dependent variables were known corresponding path coefficient could be easily assigned to each link of the paths to result in the





Table 50. Number of Possible Paths Thru Which an Independent Variable May Influence a Dependent Variable in an Eleven Variable Path Model

Dependent Variable ( $Z_j$ )	Independent Variable ( $Z_i$ )	Number of Paths Thru Which The Independent Variable Influences The Dependent Variable $Z(j-i-1)$
Z11	Z1	512
	Z2	256
	Z3	128
	Z4	64
	Z5	32
	Z6	16
	Z7	8
	Z8	4
	Z9	2
	Z10	1
Z10	Z1	256
	Z2	128
	Z3	64
	Z4	32
	Z5	16
	Z6	8
	Z7	4
	Z8	2
	Z9	1
Z9	Z1	128
	Z2	64
	Z3	32
	Z4	16
	Z5	8
	Z6	4
	Z7	2
	Z8	1
Z8	Z1	64
	Z2	32
	Z3	16
	Z4	8
	Z5	4
	Z6	2
	Z7	1
Z7	Z1	32
	Z2	16
	Z3	8
	Z4	4
	Z5	2
	Z6	1
Z6	Z1	16
	Z2	8
	Z3	4
	Z4	2
	Z5	1
Z5	Z1	8
	Z2	4
	Z3	2
	Z4	1
Z4	Z1	4
	Z2	2
	Z3	1
Z3	Z1	2
	Z2	1
Z2	Z1	1

TOTAL NO. OF POSSIBLE PATHS IN THE 11 VARIABLE MODEL = 2036



theoretically correct equations. Secondly by use of data presented in Figures 18 to 21 four other data sets for the values of the path coefficients for the calibrated 1940, 1950, 1960 and 1970 path models were prepared. These four data sets contained all the path coefficients; with zero substituted for the path coefficient corresponding with a statistically insignificant link. Finally a simple computer program was employed which multiplied the path coefficients associated with each link of each of the paths and summed the resultant partial effect coefficients for all the paths between an independent and dependent variable.

#### Results Obtained from Calculation of Total Effects of Independent Variables on Dependent Variables

Total effect coefficients of each independent variable on the dependent variables are given in Column 3 of Tables 51 to 54. The interpretation of numerical values of total effects,  $C_{j,i}$ , is that, other conditions being equal, for each standard deviation change in the independent variable the value of the dependent variable is changed by  $C_{j,i}$  standard deviation of the dependent variable. Column 4 of Tables 51 to 54 show the standard deviation of measured dependent variables. Column 6 which is obtained by multiplication of values of Column 3 by Column 4 shows the value of change in each dependent variable for one standard deviation change in the independent variable. Standard deviations of the independent variables are presented in Column 5 of these same tables to show the required amount of change in the independent variable to provide the change shown in Column 6 of the tables.



Table 51. Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1940

Dependent Variable Column 1	Independent Variable Column 2	Calculated Value of Total Effect Coefficients $(C_{j,i})$ Column 3	Standard Deviation of Measured Dependent Variable Column 4	Standard Deviation of Measured Independent Variable Column 5	Change in the Dependent Variable for One Standard Deviation Change in the Independent Variable Col. 6=Col. 3xCol. 4
RESIDENTIAL LAND USE	CITY LOCATION	0.13	477.8	61.6	62
	INTER-CITY ACCESSIBILITY	0.17		7175.4	81
	INTRA-CITY ACCESSIBILITY	0.73		299.6	349
	INDUSTRIAL LAND USE	0.23		88.7	110
	CITY POPULATION	0.95		12499.4	454
INSTITUTIONAL LAND USE	INTER-CITY ACCESSIBILITY	0.37	138.9	7175.4	51
	INTRA-CITY ACCESSIBILITY	0.36		299.6	50
	INDUSTRIAL LAND USE	0.51		88.7	45
	CITY POPULATION	0.49		12499.4	67
	COMMERCIAL LAND USE	-0.43		36.8	-60
VACANT LAND USE	INTER-CITY ACCESSIBILITY	0.18	176.1	7175.4	32
	INTRA-CITY ACCESSIBILITY	0.58		299.6	102
	INDUSTRIAL LAND USE	0.25		88.7	44
	CITY POPULATION	-0.46		12499.4	-81
	COMMERCIAL LAND USE	-0.40		36.8	-70
COMMERCIAL LAND USE	INTER-CITY ACCESSIBILITY	-0.45	36.8	7175.4	-17
	INTRA-CITY ACCESSIBILITY	0.88		299.6	32
	INDUSTRIAL LAND USE	-0.62		88.7	-23
	CITY POPULATION	1.16		12499.4	43
INDUSTRIAL LAND USE	INTER-CITY ACCESSIBILITY	0.72	88.7	7175.4	64



Table 52. Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1950

Dependent Variable Column 1	Independent Variable Column 2	Calculated Value of Total Effect Coefficients (C <sub>j,i</sub> ) Column 3	Standard Deviation of Measured Dependent Variable Column 4	Standard Deviation of Measured Independent Variable Column 5	Change in the Dependent Variable for One Standard Deviation Change in the Independent Variable Col. 6=Col. 3xCol. 4
RESIDENTIAL LAND USE	INTRA-CITY ACCESSIBILITY	0.96	512.7	237.2	492
	INDUSTRIAL LAND USE	-0.001		146.1	< 1
	EMPLOYMENT PER 1000	-0.003		31.7	- 2
	MEDIAN FAMILY INCOME	-0.004		573.3	- 2
	CITY POPULATION	0.95		12717.8	478
INSTITUTIONAL LAND USE	INTRA-CITY ACCESSIBILITY	0.84	116.4	237.2	98
	INDUSTRIAL LAND USE	-0.04		146.1	- 5
	EMPLOYMENT PER 1000	-0.06		31.7	- 7
	MEDIAN FAMILY INCOME	-0.10		573.3	-12
	CITY POPULATION	0.87		12717.8	93
VACANT LAND USE	INTRA-CITY ACCESSIBILITY	1.43	80.7	237.2	115
	INDUSTRIAL LAND USE	0.48		146.1	39
	COMMERCIAL LAND USE	0.56		64.6	45
COMMERCIAL LAND USE	INTRA-CITY ACCESSIBILITY	0.78	64.6	237.2	50
	INDUSTRIAL LAND USE	0.85		146.1	55
INDUSTRIAL LAND USE	INTER-CITY ACCESSIBILITY	0.72	146.1	7363.6	105





Table 53. Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1960

Dependent Variable Column 1	Independent Variable Column 2	Calculated Value of Total Effect Coefficients (C <sub>j,i</sub> ) Column 3	Standard Deviation of Measured Dependent Variable Column 4	Standard Deviation of Measured Independent Variable Column 5	Change in the Dependent Variable for One Standard Deviation Change in the Independent Variable Col. 6=Col.3 x Col.4
RESIDENTIAL LAND USE	CITY LOCATION	0.11	550.6	66.3	61
	INTER-CITY ACCESSIBILITY	0.21		11985.4	116
	INTRA-CITY ACCESSIBILITY	0.98		335.3	540
	INDUSTRIAL LAND USE	0.16		164.2	88
	EMPLOYMENT PER 1000	0.20		26.6	110
	CITY POPULATION	0.74		13066.1	407
	COMMERCIAL LAND USE	0.06		74.5	33
INSTITUTIONAL LAND USE	VACANT LAND USE	-0.14		270.8	-77
	INSTITUTIONAL LAND USE	-0.22		252.0	-121
VACANT LAND USE	INTER-CITY ACCESSIBILITY	0.35	252.0	11985.4	88
	INTRA-CITY ACCESSIBILITY	0.90		335.3	227
	INDUSTRIAL LAND USE	-0.73		164.2	-184
	EMPLOYMENT PER 1000	0.11		26.6	28
	CITY POPULATION	0.68		13066.1	171
	COMMERCIAL LAND USE	-0.27		74.5	-68
	VACANT LAND USE	0.61		270.8	154
COMMERCIAL LAND USE	INTER-CITY ACCESSIBILITY	0.49	270.8	11985.4	133
	INTRA-CITY ACCESSIBILITY	0.82		335.3	222
	EMPLOMENT PER 1000	-0.15		26.6	-41
INDUSTRIAL LAND USE	CITY POPULATION	-0.93		13066.1	-252
	INTER-CITY ACCESSIBILITY	0.30	74.5	11985.4	22
	INTRA-CITY ACCESSIBILITY	0.65		335.3	48
INDUSTRIAL LAND USE	INTER-CITY ACCESSIBILITY	0.42	164.2	11985.4	69
	INTRA-CITY ACCESSIBILITY	0.91		335.3	149



Table 54. Total Effects on Dependent Variables for One Standard Deviation Change in the Independent Variables - 1970

Dependent Variable Column 1	Independent Variable Column 2	Calculated Value of Total Effect Coefficients (C <sub>j,i</sub> ) Column 3	Standard Deviation of Measured Dependent Variable Column 4	Standard Deviation of Measured Independent Variable Column 5	Change in the Dependent Variable for One Standard Deviation Change in the Independent Variable Col. 6=Col.3 x Col.4
RESIDENTIAL LAND USE	CITY LOCATION	-1.34	883.7	79.7	-1184
	INTER-CITY ACCESSIBILITY	1.65		22402.2	1458
	INTRA-CITY ACCESSIBILITY	0.79		582.5	698
	INDUSTRIAL LAND USE	0.37		257.9	327
	CITY POPULATION	0.45		14395.7	398
	COMMERCIAL LAND USE	0.22		144.5	194
INSTITUTIONAL LAND USE	CITY LOCATION	-1.34	364.9	79.7	- 489
	INTER-CITY ACCESSIBILITY	1.78		22420.2	650
	INTRA-CITY ACCESSIBILITY	1.10		582.5	401
	INDUSTRIAL LAND USE	0.82		257.9	299
	EMPLOYMENT PER 1000	0.24		38.9	88
	CITY POPULATION	1.82		14395.7	664
	VACANT LAND USE	0.67		466.3	244
VACANT LAND USE	CITY LOCATION	-1.15	466.3	79.7	- 563
	INTER-CITY ACCESSIBILITY	1.35		22420.2	629
	INTRA-CITY ACCESSIBILITY	0.94		582.5	438
COMMERCIAL LAND USE	CITY LOCATION	-0.62	144.5	79.7	- 90
	INTER-CITY ACCESSIBILITY	0.88		22420.2	127
	INTRA-CITY ACCESSIBILITY	0.50		582.5	72
	INDUSTRIAL LAND USE	0.61		257.9	88
	CITY POPULATION	0.73		14395.7	106
INDUSTRIAL LAND USE	CITY POPULATION	-1.02	257.9	79.7	- 263
	INTER-CITY ACCESSIBILITY	1.45		22420.2	374
	INTRA-CITY ACCESSIBILITY	0.83		582.5	214



Results of analysis of total effects of independent variables on independent variable showed that more of the hypothesized model's variables effected the dependent variables than was evident from results of calibrations of the hypothesized model shown in Tables 41 to 44. The reason for this phenomenon is that the predictive equations of Tables 40 to 42 only contain those independent variables which directly affect the dependent variable, but the total effect analysis also considered those variables which indirectly affected the dependent variables.

Results obtained from total effect calculations are used as a tool in the sensitivity type analysis in the transportation or city planning process. Suppose an increase in a city's highway system is contemplated; then the results of total effect analysis can be used to check on the magnitude of other land uses effected by the change in the highway system. It should be noted that all the independent variables are not in the control of the planner. For example, the location of each city is fixed and the values of inter-city accessibility are only affected by the statewide planning decisions, but other independent variables can be affected by the city wide planning decisions. Qualitative comparison of the four calibrated models showed that they are different from each other. For current or future sensitivity analysis the most useful information is contained in the results of total effect analysis obtained from the 1970 model. Therefore the results obtained from 1970 total effect calculations is analyzed in the following section.



Total Effect of Independent Variables on Residential Land Use - 1970

Residential land use is negatively affected by city location and positively affected by inter-city accessibility, intra-city accessibility, industrial land use, city population and commercial land use. City location as used in the model is a dummy variable measured with reference to a horizontal line passing through Indianapolis. Therefore the implication of the negative direct and indirect effect of the city location on residential land use is that other conditions being equal 1970 sample cities in southern Indiana used less land for residential purposes than cities in northern Indiana.

The implication of total positive effect of inter-city accessibility on residential land use is that, cities with better access to other population centers have greater residential land use than cities with lesser access with other population center. It should be noted that the above statement does not mean the sample cities with greater inter-city accessibility use a greater portion of their land for residential purpose, but means that among 1970 sample cities for every increase of one standard deviation inter-city accessibility the residential land use was increased by 1.65 standard deviation of residential land use. Most cities have flexibility in enlarging the city limits, and another way of looking at the results was that the overall influence of increase in inter-city accessibility was increase in the size of city. Conversely the overall influence of city location in the southern Indiana was for the cities to remain small.





Among the variables in the control of city planner, positive effect of intra-city accessibility on residential land use was substantial, while positive effect of industrial land use, city population, and commercial land use were moderate.

#### Total Effect of Independent Variables on Institutional Land Use - 1970

Institutional land use is negatively affected by city location and positively affected by inter-city accessibility, intra-city accessibility, industrial land use, employment per 1000 persons, city population and vacant land use. Among the variables somewhat in the control of the planning process intra-city accessibility, and industrial land use had substantial influence in the institutional land use. Influence of employment rate on institutional land use is small.

#### Total Effect of Independent Variables on Vacant Land Use - 1970

City location affects vacant land use negatively; while inter-city accessibility and intra-city accessibility affect vacant land positively. Vacant land usually is the land reserved for future use therefore the implication of results obtained is that cities located in southern Indiana had the tendency to remain smaller than cities in northern Indiana; and higher accessibilities had expansionary effects on the sample cities.

#### Total Effect of Independent Variables on Commercial Land Use - 1970

City location affects commercial land use negatively; while inter-city accessibility, intra-city accessibility, industrial land use and



city population affect commercial land use positively. The implication of negative affect of city location on commercial land use was that the sample cities in southern Indiana had less land used for commercial land use.

#### Total Effect of Independent Variables on Industrial Land Use

City location affects industrial land use negatively. This implies that cities in southern Indiana had smaller industrial land use than northern cities. While inter-city accessibility and intra-city accessibility had substantial positive effect on industrial land use.

#### References

1. Campbell, D. T., and Stanley, J. C., Experimental and Quasi Experimental Design for Research, Rad McNally & Co., Chicago, 1966.
2. State Planning Services Agency, "A Directory of Indiana Local and Regional Planning and Development Commissions", Indianapolis, June 1978, p. 5.
3. U.S. Bureau of Census, U.S. Census of Population: 1970, Vol. 1, Characteristics of the Population, Part 16, Indiana, U.S.G.O.P., Washington, D. C., April 1973.
4. U.S. Bureau of Census, U.S. Census of Population: 1960, Vol. 1, Characteristics of the Population, Part 16, Indiana, U.S. Government Printing Office, Washington, D. C., 1963.
5. U.S. Bureau of Census, U.S. Census of Population: 1950, Vol. 2, Characteristics of the Population, Part 14, U.S. Government Printing Office, Washington, D. C., n.d.
6. U.S. Bureau of Census, U.S. Census of Population: 1940, Vol. II, Characteristics of the Population, Part 2, Florida-Iowa, U.S. Government Printing Office, Washington, D. C., 1943.
7. Ibid, pp. 842-844.



8. Land, Kenneth, "Principal of Path Analysis", pp. 3-37 in Edgar Borgotta (ed.), Sociological Methodology, 1969, San Francisco, Jossey Bass, p. 9.
9. Shafran, Isaac and Wegman, Frederic J., "The Influence of the Highway Network Structure on the Economic Development of West Virginia", Transportation Research Board, Record 285, 1969.
10. Hejal, S. S., Traffic Speed Report No. 88, Joint Highway Research Project, Purdue University, Lafayette, Oct. 1970.
11. Tharp, K. J., and Petty, D. F., Traffic Speed Report No. 74, Joint Highway Research Project, Purdue University, Lafayette, Nov. 1961.
12. Lang, E. R. and Michael, H. L., Traffic Speed Report No. 44, Joint Highway Research Project, Purdue University, Lafayette, May 1950.
13. Frost, R. E., "Indiana Wartime Traffic Speeds", in Proceedings of the Twenty-Ninth Annual Road School, Lafayette, Jan. 1943, pp. 99-106.
14. Vodrazka, Walter Charles, "Subclassification of the State Highway System of Indiana Based on Synthesis of Intercity Travel," Ph.D. Thesis, Purdue University, W. Lafayette, June 1968.
15. Indiana State Highway Commission, 1940 Indiana State Highway System (Scale 1 inch = 10 miles), Indianapolis, 1940.
16. Indiana State Highway Commission, 1950 State Highway System of Indiana (Scale 1 inch = 10 miles), Indianapolis, 1950.
17. Indiana State Highway Department, 1960 State Highway System of Indiana (Scale 1 inch = 10 miles), Indianapolis, 1960.
18. Indiana State Highway Commission, 1970-1971 Indiana State Highway (Scale 1 inch = 10 miles), Indianapolis, 1971.
19. Federal Writers' Project Works Progress Administration, Illinois 1939 (Highway Map), (Scale 1 inch = 12 miles), n.p., n.d.
20. "Highway Map of Illinois", Standard Highway Mileage Guide (Scale 1 inch = 16 miles), Rand McNally & Company, 1949, pp. 230-231.
21. State of Illinois, 1961 Illinois Official Highway Map, (Scale 1 inch = 12.5 miles), Springfield, 1961.
22. Illinois Division of Highways, 1970 Official Highway Map (Scale 1 inch = 12 miles), Springfield, 1970.



23. Kentucky Department of Highways, Official (1941) Kentucky Road Map, (Scale 1 inch = 12 miles), Dec. 1941, n.p.
24. Kentucky Department of Highways, Official Kentucky Road Map (Scale 1 inch = 12 miles), Frankfort, April 1950.
25. Kentucky Department of Highways, 1960 Official Kentucky Department of Highways Road Map, Frankfort, n.d.
26. Kentucky Department of Highways, Kentucky Department of Highways Map - August 1970 (Scale 1 inch = 12 miles), Frankfort, 1970.
27. Michigan State Highway Department, Official Michigan Highway Map (Scale 1 inch = 14.5 miles), Winter Edition 1939-1940, 1939, Lansing.
28. Michigan State Highway Department, Michigan Official Highway Map, 1950 (Scale 1 inch = 14.5 miles), Lansing, April 1950.
29. Michigan State Highway Department, Michigan 1960 Official Highway Map (Scale 1 inch = 14.5 miles), Lansing.
30. Michigan State Highway Commission, Michigan 1970 Official Highway Map (Scale 1 inch = 14.5 miles), Lansing, n.d.
31. U.S. Geological Survey, "The Federal Aid Highway System Program Map," 1 inch = 7.9 miles, map 1 (northern Ohio) and Map 2 (southern Ohio), n.p., June 1, 1939.
32. Ohio Department of Highways, Ohio Official Highway Map, 1959-1960, (Scale 1 inch = 10 miles), n.d., n.p.
33. Ohio Department of Highways, Ohio Official Highway Maps, 1950-1960, (Scale 1 inch = 10 miles), n.p., 1960.
34. Ohio Department of Highways, 1969 Official Ohio Highway Map and Economic Digest, (Scale 1 inch = 10 miles), Columbus, n.d.
35. Vodrazka, Walter Charles, Ibid., p. 92.
36. Homburger, Wolfgang S., "Traffic Estimation Computer Programs for Educational Purposes - 3rd Edition," Institute of Transportation and Traffic Engineering, University of California, Berkeley, Oct. 1970, pp. 31-35.
37. Ibid, pp. 40-45.
38. "ASCS Aerial Photography Status Maps as of December 31, 1974", Agricultural Stabilization and Conservation Service, U.S. Department of Agriculture, Salt Lake City, Utah.





39. Miles, Robert D., "Availability of Aerial Photographs", in Standards of Practice for Indiana Society of Professional Land Surveyors, sec. VIII, Item 6, pp. 13-20, n.d.
40. Davis, Jeane M., "Uses of Airphotos for Rural and Urban Planning", Agricultural Handbook No. 315, U.S. Department of Agriculture, June 1966, p. 9.
41. Hadfield, Sam M. and Orzeske, John D., "A Sampling Technique for Updating a Quantitative Land Use Inventory", Transportation Research Board, Record 126, 1966, p. 81.
42. Falkner, E., "Land Use Changes in the Parkway School District", Photogrammetric Engineering, XXXIV, Jan., 1968, pp. 52-57.
43. U.S. Department of Agriculture, "Sampling, Coding and Storing Flood Plain Data", Agriculture Handbook No. 237, August 1962.
44. Ibid.
45. Branch, Melville C., City Planning and Aerial Information, Harvard University Press, Cambridge, 1971.
46. Manual of Photographic Interpretation, American Society of Photogrammetry, 1960.
47. Mgbejume, Joel Keboku, "Application of the Techniques of Aerial Photographic Interpretation in the Analysis of Land Use Changes in GreenTreeBorough, Pennsylvania", Masters Thesis, University of Pittsburgh, Pittsburgh, 1962.
48. Pownall, L. L., "Airphoto Interpretation of Urban Use in Madison, Wisconsin", Photogrammetric Engineering, 1950, pp. 414-426.
49. Collins, W. Gordon, et al, "The Acquisition of Urban Land Use Information from Aerial Photographs of the City of Leeds (Great Britain)", Photographic Engineering, Vol. 27, 1971, pp. 71-92.
50. Letter from Meyer Zitter, Chief Population Division, Bureau of Census, to Hadi Yamin, Purdue University, dated Sept. 27, 1977.
51. Hutchinson, B. G., Principles of Urban Transport Systems Planning, Scripta Book Company, Washington, D. C., 1974, p. 58.
52. Draper, N. R. and Smith, H., Applied Regression Analysis, John Wiley & Sons, Inc., New York, 1966, pp. 167-169.
53. Nie, Norman, et al, Statistical Package for the Social Sciences, McGraw-Hill Book Company, New York, 1975.



54. Draper, N. R. and Smith, H., pp. 21-24 and pp. 121-122.
55. "Confidence and Prediction Intervals for Y-Program Name: CPIYL", maintained by Purdue University Statistics Department, Nov., 1974.
56. "SPSS Regression Reference, CS SPSS 15", Purdue University Computer Center, Nov. 1976, p. 16.



## CHAPTER FIVE

## CONCLUSION

A path model was hypothesized to explain the relationship of transportation development and city development for non-SMSA Indiana cities whose population was less than 50,000.

1. Data from 1940, 1950, 1960 and 1970 were used to calibrate the path model for the four decades. The results of the calibrated models which were presented in Tables 41 to 44 as well as Figures 18 to 21 are summarized in Table 55. This table shows the percent of the variations of land uses explained by the twenty equations (five land uses in four time periods). Six of these equations explained more than 90 percent of land use variation measured by them. Nine of the equations explained 73 to 88 percent of the measured variations; while the remaining five equations measured 44 to 54 percent of measured land use variations.
2. Three cities, one from each city size classification, were removed from each of the four data sets and the four path models were recalibrated with the use of the remainder of data. It was found that recalibrated models could estimate the land uses for the cities not used in their calibration with a ninety-five percent significant level.



Table 55. Summary of Results Obtained from Calibration of 1940, 1950, 1960 and 1970 Path Models

Independent Variables	Percent of Variations Explained by the Path Model ( $R^2$ )			
	1940 Model	1950 Model	1960 Model	1970 Model
Residential Land Use	93.44	95.40	98.26	97.70
Institutional Land Use	77.47	75.89	96.50	94.13
Vacant Land Use	74.87*	44.22	86.93	82.29
Commercial Land Use	49.10	72.97	41.79	54.45
Industrial Land Use	52.31	82.87	82.26	88.52

Note:  $\alpha$  = (1-significance level) for all equations is 0.00 except as noted by (\*) where  $\alpha$  = 0.03.





3. The four reduced path models obtained from calibration of the hypothesized transportation-city development path model were qualitatively compared and it was determined the four models are generally different, and it was not feasible to combine the data for the four decades to calibrate an overall model.
4. Total effects coefficients were calculated for the four models. These coefficients gave the total change, in standard deviations, of dependent variables for one standard deviation change in each independent variable.

The results obtained from this research are useful as transportation and city planning tools for those non-SMSA cities with population greater than 2,500 and less than 50,000.

1. Although the calibrated path models do not include time as a variable it is hoped that the calibrated 1970 model can be used to estimate land use changes as consequences of future planning decisions. To test the feasibility of this procedure; the five land use variables for the 24 sample cities measured in 1970 were estimated with the use of 1960 model. The results of this test which is presented in Appendix C-1 to C-5 showed that from a total of 120 observations; 106 were within their corresponding 95 percent confidence interval as estimated by the 1960 model.
2. Total effect coefficients obtained from the 1970 calibrated path model are a powerful planning tool because the total effect of a proposed change in a dependent variable (while controlling for all other variables) is given by the total



effect coefficient. The simplicity of use of total effect coefficients may make them very useful in environmental impact studies of the proposed transportation or industrial projects.

3. This research demonstrated applicability of cross sectional analysis in conjunction with the statistical path model to explain the relationship of transportation development to urban development. It is believed the approach in the long run is more fruitful than the traditional before and after studies.

It is recommended that similar research in other geographical areas be carried out to enlarge the inference space of the experiment. It would be helpful for the better understanding of the relationship of transportation and city development to compare the results of similar model development for other states with the results obtained in this research.

Two areas of improvement are anticipated for future research in the relationship of transportation and city development:

1. In the current project data was collected for the four different periods of 1940, 1950, 1960, and 1970. The transportation-city development path model was calibrated for each of the four periods and it was found that in all the four periods transportation development had a significant effect in the city development, but qualitatively it was decided that the four reduced models were different from each other. The four calibrated models could not be quantitatively compared because there were only about 24 sample cities in each



period. One possible way to quantitatively compare the four models was to accept time and time's interaction with the 11 hypothesized variables as new variables and test for significance of time and time interactions. But then there would be 23 variables in the model and there were not sufficient data points to test the statistical significance of the variables. It is recommended that enough sample points be chosen to test for time and its interaction or if the project resources do not allow sufficient sample points for more than one time period to concentrate on the latest period for which data can be collected. For example, in this research the most useful part of the research was the effort used on analysis of the 1970 data.

2. It is hoped that future research improves on the operational definition of intra-city accessibility as used in this project (acres of land used for highways). The usual definition of intra-city accessibility,

$$acc_i = \sum_{J=1}^n a_J f_{iJ}$$

as discussed in page 94 is not conveniently calculated for this type of project; where the city as a whole is the sample point and information about inter-zonal travel is not readily available.



## LIST OF REFERENCES





## LIST OF REFERENCES

Adkins, William G., "Effects of the Dallas Central Expressway on Land Values and Land Use", Texas Transportation Institute, Sept. 1957.

"ASCS Aerial Photography Status Maps as of December 31, 1974", Agricultural Stabilization and Conservation Service, U.S. Department of Agriculture, Salt Lake City, Utah.

Babcock, W. F., "An Analysis of the Impacts of Freeways on Urban Land Developments in North Carolina (Including Guidelines for Highway Planning and Design)", North Carolina State University, June 1974.

Bascom, S. E., et al, Secondary Impacts of Transportation and Wastewater Investments: Review and Bibliography, U.S. Environmental Protection Agency, Report No. EPA 600/5-75-002, January 1975.

Bone, A. J. and Wohl, M., "Massachusetts Route 128 Impact Study", Highway Research Board, Bulletin 227, 1959.

Branch, Melville C., City Planning and Aerial Information, Harvard University Press, Cambridge, 1971.

Branham, A. K., et al, "Economic Evaluation of Two Indiana Bypasses", Highway Research Board, Bulletin 67, 1953.

Brown, F. A., and Michal, H. L., "The Impact on Land Value of a Major Highway Interchange Near a Metropolitan Area", Joint Highway Research Project, No. 34, December 1973.

Bureau of Public Roads, Highways and Economic and Social Changes, 1964.

Campbell, D. T., and Stanley, J. C., Experimental and Quasi Experimental Design for Research, Rand McNally & Co., Chicago, 1966.

Charles River Associates, Inc., Measurement of the Effects of Transportation Changes, prepared for DOT, Sept. 1972.

"Confidence and Prediction Intervals for Y-Program Name: CPIYL", Purdue University Statistics Department, W. Lafayette, n.d.

Collins, W. Gordon, et al, "The Acquisition of Urban Land Use Information from Aerial Photographs of the City of Leeds (Great Britain)", Photographic Engineering, Vol. 27, 1971, pp. 71-92.



Davis, Jeane M., "Uses of Airphotos for Rural and Urban Planning", Agricultural Handbook No. 315, U.S. Department of Agriculture, June 1966.

Draper, N. R. and Smith, H., Applied Regression Analysis, John Wiley & Sons, Inc., New York, 1966.

Evans, Eugene Glem, "Impact of Kokomo By-Pass from 1950 to 1964", M.S. Thesis, Purdue University, August 1965.

Fabbroni, Lawrence P., "Land Use Development at Interstate Interchanges in Indiana", Joint Highway Research Project, No. 9, May 1973.

Falkner, E., "Land Use Changes in the Parkway School District", Photogrammetric Engineering, XXXIV, Jan. 1968, pp. 52-57.

Federal Highway Administration, Economic and Social Effect of Highways, 1972.

Federal Highway Administration, Economic and Social Effect of Highways, 1974.

Federal Highway Administration, Social and Economic Effects of Highways, 1976.

Federal Writers' Project Works Progress Administration, Illinois 1939 (Highway Map), (Scale 1 inch = 12 miles), n.p., n.d.

Fletcher, Joseph A., "Early Effects of Rural Interstate Highways", M.S. Thesis, Purdue University, June 1961.

Frost, R. E., "Indiana Wartime Traffic Speeds", in Proceedings of the Twenty-Ninth Annual Road School, Lafayette, Jan. 1943.

Frye, Frederick, "Alternative Multimodal Passenger Transportation System: Comparative Economic Analysis", National Cooperative Highway Research Program, Report 146, 1973.

Hadfield, Sam M. and Orzeske, John D., "A Sampling Technique for Updating a Quantitative Land Use Inventory", Transportation Research Board, Record 126, 1966.

Hammerschlag, Dieter, et al, "The Interstate Highway System and Urban Structure: A Force for Change in Rhode Island", University of Rhode Island, July 1976.

Harrison, Joseph W., "Economic Studies Progress Report No. 2: Study of the Economic Effects of the U.S. Route 11 By-Pass at Lexington, Virginia on Business Volume and Composition", Virginia Council of Highway Investigation and Research, Charlottesville, Virginia, 1958.



Hejal, S. S., Traffic Speed Report No. 88, Joint Highway Research Project, Purdue University, Lafayette, Oct. 1970.

"Highway Map of Illinois", Standard Highway Mileage Guide (Scale 1 inch = 16 miles), Rand McNally & Company, 1949, pp. 230-231.

Homburger, Wolfgang S., "Traffic Estimation Computer Programs for Educational Purposes - 3rd Edition", Institute of Transportation and Traffic Engineering, University of California, Berkeley, Oct. 1970.

Horwood, Edgar M. and Boyce, Ronal R., "Measurement of Central Business District Change and Urban Highway Impact", Highway Economic Studies, University of Washington, 1959.

Hutchinson, B. G., Principle of Urban Transport System Planning, McGraw-Hill, New York, 1974.

Illinois Division of Highways, 1970 Official Highway Map (Scale 1 inch = 12 miles), Springfield, 1970.

Indiana State Highway Commission, 1940 Indiana State Highway System (Scale 1 inch = 10 miles), Indianapolis, 1940.

Indiana State Highway Commission, 1950 State Highway System of Indiana (Scale 1 inch = 10 miles), Indianapolis, 1950.

Indiana State Highway Department, 1960 State Highway System of Indiana (Scale 1 inch = 10 miles), Indianapolis, 1960.

Indiana State Highway Commission, 1970-1971 Indiana State Highway (Scale 1 inch = 10 miles), Indianapolis, 1971.

Isibor, Edward I., "Modelling the Impact of Highway Improvements on the Value of Adjacent Land Parcels", Ph.D. Thesis, Purdue University, January 1970.

Kelly, J. F., "Tulare By-Pass", Reprint from California Highway and Public Work, May-June 1956.

Kentucky Department of Highways, Official (1941) Kentucky Road Map, (Scale 1 inch = 12 miles), Dec. 1941, n.p.

Kentucky Department of Highways, Official Kentucky Road Map (Scale 1 inch = 12 miles), Frankfort, April 1950.

Kentucky Department of Highways, 1960 Official Kentucky Department of Highways Road Map, Frankfort, n.d.

Kentucky Department of Highways, Kentucky Department of Highways Map - August 1970 (Scale 1 inch = 12 miles), Frankfort, 1970.



Kingham, Ian R., "Suburban Highways and Roads as Instruments of Land Use Change", Transportation Research Board, Record 565, 1976.

Land, Kenneth, "Principles of Path Analysis", pp. 3-37 in Edgar Borgetta (ed.), Sociological Methodology, Jossey Bass, San Francisco, 1969.

Lang, E. R. and Michael, H. L., Traffic Speed Report No. 44, Joint Highway Research Project, Purdue University, Lafayette, May 1950.

Letter from Meyer Zitter, Chief Population Division, Bureau of Census, to Hadi Yamin, Purdue University, dated Sept. 27, 1977.

Lohr, Alan F., "Early Impact of a Highway Improvement on an Urban Area", M.S. Thesis, Purdue University, 1962.

Manheim, Marvin L., et al, "Transportation Decision-Making: A Guide to Social and Environmental Considerations", National Cooperative Highway Research Program, Report 156, 1975.

Manual of Photographic Interpretation, American Society of Photogrammetry, 1960.

Michael, Harold L., "An Economic Evaluation of the Kokomo Indiana Bypass", M.S. Thesis, Purdue University, August 1951.

Michigan State Highway Department, Official Michigan Highway Map (Scale 1 inch = 14.5 miles), Winter Edition 1939-1940, 1939, Lansing.

Michigan State Highway Department, Michigan Official Highway Map, 1950 (Scale 1 inch = 14.5 miles), Lansing, April 1950.

Michigan State Highway Department, Michigan 1960 Official Highway Map (Scale 1 inch = 14.5 miles), Lansing.

Michigan State Highway Commission, Michigan 1970 Official Highway Map (Scale 1 inch = 14.5 miles), Lansing, n.d.

Mgbejume, Joel Keboku, "Application of the Techniques of Aerial Photographic Interpretation in the Analysis of Land Use Changes in GreenTree Borough, Pennsylvania", Masters Thesis, University of Pittsburgh, Pittsburgh, 1962.

Miles, Robert D., "Availability of Aerial Photographs", in Standards of Practice for Indiana Society of Professional Land Surveyors, sec. VIII, Item 6, pp. 13-20, n.d.

Moore, T. W., et al, An Introduction to Urban Development Models and Guidelines for Their Use in Urban Transportation Planning, U.S. Department of Transportation, Oct. 1975.





Nie, Norman, et al, Statistical Package for the Social Sciences, SPSS, Second Edition, McGraw-Hill Book Company, New York, 1975.

Ohio Department of Highways, Ohio Official Highway Map, 1959-1960, (Scale 1 inch = 10 miles), n.d., n.p.

Ohio Department of Highways, Ohio Official Highway Maps, 1950-1960, (Scale 1 inch = 10 miles), n.p., 1960.

Ohio Department of Highways, 1969 Official Ohio Highway Map and Economic Digest, (Scale 1 inch = 10 miles), Columbus, n.d.

Pownall, L. L., "Airphoto Interpretation of Urban Use in Madison, Wisconsin", Photogrammetric Engineering, 1950, pp. 414-426.

Putman, Stephen H., The Interrelationship of Transportation Development and Land Development, prepared for Federal Highway Administration, Vol. I and II, 1973.

Shafran, Issac and Wegman, Frederic J., "The Influence of the Highway Network Structure on the Economic Development of West Virginia", Transportation Research Board, Record 285, 1969.

Skorpa, Liverd, et al, Transportation Impact Studies: A Review with Emphasis on Rural Areas, Texas University at Austin, NTIS PB-239463, 1974.

SPSS Regression Reference, GS SPSS 15", Purdue University Computer Center, W. Lafayette, Nov. 1976.

State of Illinois, 1961 Illinois Official Highway Map, (Scale 1 inch = 12.5 miles), Springfield, 1961.

Tharp, K. J., and Petty, D. F., Traffic Speed Report No. 74, Joint Highway Research Project, Purdue University, Lafayette, Nov. 1961.

Texas Transportation Institute, "Economic Impact Study - Waxahachie, Texas", Bulletin No. 35, March 1966.

State Planning Services Agency, "A Directory of Indiana Local and Regional Planning and Development Commissions", Indianapolis, June 1978.

U.S. Bureau of Census, U.S. Census of Population: 1940, Vol. II, Characteristics of the Population; Part 2 Florida-Iowa, U.S. Government Printing Office, Washington, D. C., 1943.

U.S. Bureau of Census, U.S. Census of Population: 1950, Vol. 2, Characteristics of the Population Part 14, U.S. Government Printing Office, Washington, D. C., n.d.



U.S. Bureau of Census, U.S. Census of Population: 1960, Vol. I, Characteristics of the Population, Part 16, Indiana, U.S. Government Printing Office, Washington, D. C. 1963.

U.S. Bureau of Census, U.S. Census of Population: 1970, Vol. 1, Characteristics of the Population, Part 16, Indiana, U.S. Government Printing Office, Washington, D. C., April 1973.

U.S. Department of Agriculture, "Sampling, Coding and Storing Flood Plain Data", Agriculture Handbook No. 237, August 1962.

U.S. Geological Survey, The Federal Aid Highway System Program Map, 1 inch = 7.9 miles, map 1 (northern Ohio) and Map 2 (Southern Ohio), n.p., June 1, 1939.

Vodrazka, Walter Charles, "Subclassification of the State Highway System of Indiana Based on Synthesis of Intercity Travel", Ph.D. Thesis, Purdue University, W. Lafayette, June 1968.

Vargha, Louis A., "Highway Bypasses, Natural Barriers and Growth in Michigan", Highway Research Board, Bulletin, 250, 1960.

Wold, Herman, "Causal Inference from Observational Data: A Review of Ends and Means", Journal of Royal Statistical Society, Sec. A., 1956, pp. 28-50.



## APPENDICES

The Appendices are not included in this copy of the Report. The three (3) Appendices are as follows:

Appendix A - Program Used to Calculate Interaction Potential and Interaction Potential Tables for 1940, 1950, 1960 and 1970	pp. 176-208
Appendix B - Set of All Possible Paths from Each Independent Variable to Each Dependent Variable	pp. 209-238
Appendix C - Tables of Estimated Land Uses from the 1960 Model	pp. 239-243

A copy of each of the above Appendices may be obtained at the cost of duplication by inquiry to:

Joint Highway Research Project  
Civil Engineering Building  
Purdue University  
West Lafayette, IN 47907





